View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Applied Energy 177 (2016) 692–702

Contents lists available at ScienceDirect

Applied Energy

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

Understanding electricity consumption: A comparative contribution of building factors, socio-demographics, appliances, behaviours and attitudes

Gesche Huebner [a,*], David Shipworth [a], Ian Hamilton [a], Zaid Chalabi [b], Tadj Oreszczyn [a]

[a] UCL Energy Institute, 14 Upper Woburn Place, London WC1H 0AE, England, United Kingdom
[b] London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, England, United Kingdom

HIGHLIGHTS

- Multicollinearity is an issue in analysing electricity consumption data.
- Appliance ownership and use are most important in understanding electricity consumption.
- Dwelling and household size are likewise significant predictors.
- Reported attitudes hardly play a role.

ARTICLE INFO

Article history:
Received 28 November 2015
Received in revised form 12 April 2016
Accepted 14 April 2016
Available online 3 June 2016

Keywords:
Electricity demand
Residential buildings
Socio-demographics
Appliances
Lasso regression
Multicollinearity
English homes
Gas-central heating

ABSTRACT

This paper tests to what extent different types of variables (building factors, socio-demographics, appliance ownership and use, attitudes and self-reported behaviours) explain annualized electricity consumption in residential buildings with gas-fuelled space and water heating. It then shows which individual variables have the highest explanatory power. In contrast to many other studies, the study recognizes the problem of multicollinearity between predictors in regression analysis and uses Lasso regression to address this issue.

Using data from a sample of 845 English households collected in 2011/12, a comparison of four separate regression models showed that a model with the predictors of appliance ownership and use, including lighting, explained the largest share, 34%, of variability in electricity consumption. Socio-demographic variables on their own explained about 21% of the variability in electricity consumption with household size the most important predictor. Building variables only played a small role, presumably because heating energy consumption is not included, with only building size being a significant predictor. Self-reported energy-related behaviour, opinions about climate change and 'green lifestyle' were negligible. A combined model, encompassing all predictors, explained only 39% of all variability (adjusted R^2 = 34%), i.e. adding little above an appliance and lighting model only. Appliance variables together with household size and dwelling size were consistently significant predictors of energy consumption.

The study highlights that when attempting to explain English household non-heating electricity consumption that variables directly influenced by people in the household have the strongest predictive power when taken together.

© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Throughout the OECD (Organisation for Economic Co-operation and Development) countries, residential dwellings are responsible for a large share of national carbon emissions, with the UK being typical at around 25% [1]. Given that heating accounts for the greatest part of residential energy use and the high prevalence of natural gas-fired heating systems in the UK, gas consumption is substantially higher than electricity consumption [2]. Despite energy efficiency improvements in electrical appliances over the last 40 years, electricity consumption of domestic appliances has increased by about 2% per year over this period whereas it has fallen slightly overall [2], making electricity consumption an

* Corresponding author.
E-mail addresses: g.huebner@ucl.ac.uk (G. Huebner), d.shipworth@ucl.ac.uk (D. Shipworth), i.hamilton@ucl.ac.uk (I. Hamilton), Zaid.Chalabi@lshtm.ac.uk (Z. Chalabi), t.oreszczyn@ucl.ac.uk (T. Oreszczyn).

http://dx.doi.org/10.1016/j.apenergy.2016.04.075
0306-2619/© 2016 The Authors. Published by Elsevier Ltd.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

important research area. In 2014, consumer electronics (e.g. TVs mobile phones) were the largest consuming group of domestic appliances with an estimated consumption of 1.8 Mtoe, followed by wet appliances (e.g. electric showers) with 1.3 Mtoe, cooking (1.1 Mtoe), cold appliances (1.1 Mtoe) and lighting (1.0 Mtoe) [2]. Hence, given the need to reduce carbon emissions significantly to mitigate climate change and meet legal targets [3], it is important to understand what factors explain residential electricity consumption and how to minimise or reduce it.

A recent paper has shown that total energy consumption (most of which will end up as heat in the building) in English households is largely explained by dwelling characteristics [4], with a comparatively small contribution of socio-demographics, self-reported behaviours, and attitudes towards environmentally significant behaviour and climate change. However, for electricity consumption without space and water heating it is expected that appliances ownership and use and socio-demographics would have a bigger impact [5].

1.1. Literature review of previous findings on determinants of electricity consumption

We only review empirically based studies, not modelling studies (for those, see e.g. [6,7] for reasons of methodological comparability.

1.1.1. Impact of building characteristics

Regarding the likely impact of dwelling characteristics on electrical energy consumption, one important factor is the composition of the sample. Given the well-documented effect of building characteristics on energy consumption when including heating (e.g. [4,8,9]), the impact of dwelling characteristics can be expected to vary in magnitude depending on whether space heating is included in the electricity consumption.

Homes using electricity for heating (e.g. [10–13]) and hot water are shown to use more electricity compared to those heating water with gas [12,13]. Another rather self-evident effect is that in geographically diverse sample of homes location plays a role (e.g. [14–16]) due to differential demand for heating (if included) and cooling.

Regarding the impact of building variables, a common finding was that detached houses have been reported to have highest electricity consumption, both when controlling for other variables and when not (Wyatt [17], Brounen et al. [5], Yohanis et al. [18]). In one study, the effect of building type only played a role in winter data when heating loads contributed significantly to electricity consumption [15].

Regarding building age, results vary across studies, with some studies finding a non-linear relationship between building age and electricity consumption (e.g. Brounen et al. [5], Wyatt [17]), others finding an effect only in subgroups of homes with electric heating but not gas heating [10], and others reporting no effect [15].

A larger floor area is generally associated with greater electricity use [5,15,16,18], and a higher floor area is more likely to signify a high consumption household [13].

Regarding the impact of additional numbers of rooms, results varied, partly depending on whether floor size was used as an additional predictor: Once controlling for floor area, Brounen et al. [5] found a negative effect of additional numbers of rooms on electricity consumption; not controlling for floor area, Tiwari [19] found a positive effect. Wiesmann et al. [16] found no effect of number of rooms; Bedir et al. [20] found a negative effect of additional bedrooms but a positive effect for additional study and hobbies room, potentially because the latter ones are associated with additional appliances.

Number of storeys, insulation of external walls, insulation of loft, and energy saving windows had no significant effect on annual electricity consumption [10], in that sample, only a subset of homes used electricity for heating, which might explain why these factors played no role.

To summarize, the effect of building variables highly depends on what electricity is used for, i.e. whether it includes space and water heating. General findings are that greater floor area and detached houses use more electricity.

1.1.2. Impact of socio-demographic variables

A larger household size is generally associated with higher electricity use; however, the effect is not necessarily shown to be linear and depends on how the variable is coded. Using household size as continuous predictor showed that a larger household was associated with greater electricity consumption [10,12,20,21]. However, other papers report that whilst larger households use more electricity, the per-capita consumption is lower and hence coded household size as a categorical predictor [15,16,18,22]. Looking at what factors define being a high electricity user; Jones and Lomas [13] found that households with three or more occupants were more likely to be high consumers than homes with one or two occupants. They also found that households with teenagers were more likely to be high consumers of electricity, as did Brounen et al. [5].

Regarding the effect of age of householders, results were ambiguous. with some studies finding a non-linear effect (e.g. [15], others reporting no effect [20]. A higher use with older head of household was reported by Tiwari [19], and similarly, lower consumption if the head of household was younger than 45 years [14]. Regarding the probability of being a high consumer, Jones and Lomas [13] found that dwellings with a head of household over 65 years old were significantly less likely to be high electricity consumers than those between 36 and 50 years old; other age categories did not differ significantly from this reference category.

Income is another much studied variable, with several studies finding that households with higher income were more likely to be in the category of high consumers of electricity [13], or consumed more electricity, respectively [5,16,18,19,21,22], even though Wiesmann [16] stated the effect was relatively small once other variables were controlled for. However, Kavousian et al. [15] found no relationship between income and electricity consumption, and suggest that this might be because the income effect is mediated by appliance ownership which was a separate variable in the analysis. Bedir et al. [20] report that whilst income on its own is related to electricity consumption, it is not a significant predictor once controlling for other variables in an regression analysis, including appliances, lending support to Kavousian's findings.

Hence, both the statistical significance of an income effect and its strength might depend on what other variables are controlled for in the analysis.

1.1.3. Impact of appliances and lighting

Appliance ownership was as an explanatory variable included in several studies, with a general finding that owning more appliances and/or using them for longer is associated with greater electricity consumption [16,20]. This association with greater electricity consumption also holds when taking power consumption of appliances into account [19], and when relating base load to overall electricity consumption [18]. Specific appliances associated with greater electricity consumption were the number of refrigerators and entertainment devices for the daily minimum electricity consumption, and electric water heater, electric clothes dryer, and Spas/Pools for the daily maximum consumption [15]. Bedir et al. [20] reported that general use appliances and hobby appliances use were significant predictors (when controlling for household and dwelling characteristics) but that food preparation and cleaning appliances were not. In addition, number of showers per week and number of dryer loads per week were significant predictors [20], along with presence of an air-conditioning unit [12,21]. The number of halogen or energy-saving light bulbs however were not significant predictors [20].

Regarding cooking appliances, the number of electric stoves [15], and electric cooking vs. other cooking fuels [23] has been linked to higher electricity consumption.

1.1.4. Impact of other occupant variables

Under 'other occupant variables', we considered variables such as climate change concern, self-reported energy saving actions, and attitudes towards being green. A rationale for including 'other occupant variables' as explanatory variables in the regression model is that behaviour change interventions can lead to significant reductions of electricity consumption (for a review, see [24] and they operate through such variables. Results, however, are ambiguous, for example Darby et al. [25] described electricity reductions between 5% and 15% for interested users of in-home displays showing both continuous and historic usage, whereas a recent Swedish study found no effect of in-home-displays [26]. This section is not concerned with self-reported levels of environmental concern and pro-environmental behaviour per se (see e.g. [22]) but with studies that linked those variables to empirical electricity consumption or energy consumption in general. Vringer and Blok [27] found no evidence of a relationship between domestic energy requirements and values including problem perceptions of climate change. Abrahamse and Steg [28] reported that psychological variables such as attitudes and perceived behavioural control were not related to energy consumption but only to energy savings in an intervention program. Similarly Brandon and Lewis [29] found that environmental attitudes did not predict historic energy consumption but were related to energy savings in a subsequent intervention program. Huebner et al. [30] evidenced that self-reported habit strength was significantly related to self-reported energy consuming behaviours and to actual energy consumption (combined gas and electricity), when controlling for several building factors. However, the latter sample was small and restricted to social housing tenants only, making it unclear to what extent results can be generalized.

One study found that households that have expressed a motivation to buy energy-efficient appliances and air conditioners have higher levels of daily minimum consumption, contrary to what might be expected [15]. The same authors found that occupants who reported turning off lights when not in use had higher electricity consumption, contrary to what might be expected.

The only relationship between environmental concern variables and electricity consumption was found by Cramer et al. [21] who found that whilst there was no direct significant effect of the environmentalism and conservation scale on electricity consumption, there was a small effect of the environmentalism scale on the appliance index, indicating that those with greater ecological concern have or use fewer (energy-intensive) appliances.

Overall, there is little evidence of a relationship between 'other occupant variables' and actual electricity consumption.

1.2. Aim and scope of this paper

The aim of the research presented in this paper is to grow our understanding of what determines electricity consumption in homes, with the emphasis on the contribution of various classes of predictors.

As the literature review above shows, for many factors, findings are somewhat ambiguous and depend on inclusion of control variables and definition of electricity consumption. Hence, comparison across studies is difficult, in particular, because there is no study that looked at a large number of potential variables of interest. The only other studies that come close were performed by Bedir et al. [20] and Cramer et al. [21]. Bedir constructed consecutive regression models, starting with appliance data, and adding household and dwelling characteristics (but no other occupant variables) but with a sample of about 320 households from the districts only in the Netherlands. Cramer's sample was limited to less than 200 households and looked at summer consumption only. No individual appliances were considered, but an appliance index was created, making it a challenge to disentangle the specific effect of appliance use. In contrast, the analysis in this paper is based on a nationally representative sample of 845 households with variables on a large number building characteristics, socio-demographics, appliance use, and 'other occupant variables'. Hence, the relatively large dataset allows us to uniquely quantify the impact of various types of predictors within the same sample.

Furthermore, the majority of studies reviewed did not report checking and controlling for multicollinearity. Multicollinearity occurs when two or more predictor variables in a multiple regression model are highly correlated. Examples include the likelihood of higher income correlating with more appliances, and flats tending to be smaller than detached houses, introducing a confounding between dwelling type and size. The presence of multicollinearity means that regression coefficients cannot be reliably interpreted. In our study, for each regression analysis, variance inflation factors are inspected to see if multicollinearity exists, and if it does, Lasso regression is carried out which sets redundant predictors to zero [31], therefore performing variable selection and removing multicollinearity.

Finally, our paper looks at determinants of electricity consumption in dwellings with gas central space and water heating as their main heating system only; hence, heating could only impact on electricity consumption if householders used supplementary electric heating (which was controlled for in the analysis as this information was available). We first perform regression models of the four individual predictor classes (building variables, socio-demographics, lighting/appliance variables, and other occupant variables), checking and if necessary controlling for multicollinearity. In a final step, we combine all variables together in one regression analysis.

2. Methods

2.1. Data set

The data analysed for this paper were collected as part of the Energy Follow-Up Survey (EFUS) commissioned by the Department of Energy and Climate Change [32]. Householders were asked in an interview survey about a wide range of issues, such as details of their dwelling and opinions about climate change. Based on gas and electricity meter readings obtained in a subsample of homes, annualised energy consumption was estimated. In this paper only annual electricity consumption was considered. All households in the survey had also participated in the English Housing Survey (EHS) that collects detailed information about the English building stock. The sample size for EFUS was $N = 2616$; meter readings were available for $N = 1345$ households. Of those 1345 households, another 500 were excluded from the sample based on the following five criteria:

(1) There was a positive reply to the question if physical changes to the dwelling had been carried out since the last EHS and/or to the question if the household composition had been changed since the last EHS; as it was not recorded what exactly changed and when, the impact on electricity consumption could not be assessed (345 cases excluded).
(2) A main heating system other than gas central heating. This was done to avoid too small subsamples for e.g. heating with wood, and to avoid the total energy consumption being dominated by heating which would be the case for electrically heated homes (50 cases excluded).
(3) Hot water system not linked to the gas central heating system (22 cases excluded), again, subsamples would be too small for meaningful analysis.
(4) The log-transformed annual electricity consumption was considered an extreme value, i.e. ±3 SD from the sample mean of energy consumption (12 cases excluded).
(5) Missing data on the attitudinal variables which would have made it necessary to code the variable as categorical instead of using them as a continuous predictor and creating a non-informative category of "missing data" (55 cases excluded).
(6) Households that answered questions about lighting in either main bedroom, kitchen, or living room with "not applicable" or "not lit" (16 cases excluded). These categories were too small to be used in analysis, were hard to interpret, and additionally created many missing data in other variables (such as energy-efficient lighting).

Hence the total remaining sample size was $N = 845$ households which formed the basis for all the analyses carried out in this paper.

2.2. Independent variables

This section explains the variables used as predictors in subsequent regression analyses. Table 1 shows the building variables used and their frequencies or summary statistics (M means Mean, SD means standard deviation for the continuous variables). Note that in Tables 1, 2 and 4 the category printed in bold indicates the reference category for later regression analyses. The abbreviation in parentheses after variable name indicates how the variable is abbreviated in reporting of results.

Because only households with gas central heating were considered which also had water heating coupled to the gas central heating system there are no variables on the main heating system.

Table 2 shows descriptive information for the socio-demographic variables. Income was coded as equivalized income, i.e. household incomes were adjusted for household composition and size such that those incomes can be directly compared with each other. This means increasing incomes of small households and decreasing the incomes of large households. The extent of these increases and decreases is determined by an internationally agreed set of scales. Equivalized income was chosen as it is considered to provide a better indication of household disposable income which might in turn be predictor of expenditure on electricity consuming appliances as well as financial pressure on electricity bills.

Table 3 summarizes the continuous variables representing data on appliances and lighting, and Table 4 the categorical variables. Note that both mean and standard deviation, and median and interquartile range are stated. Most continuous variables are not normally distributed but were positively skewed; hence, the mean is not the most informative descriptive statistic.

For the three items indicating weekly appliance use (laundry, dishwasher, tumble dryer) and for the hours of TV watched per day, a null (zero) value was assigned if the appliance was not owned (and of course, when not using it at all). This was done to be able to use the variable as a continuous one; otherwise, a categorical variable would have been needed to code for not having a certain appliance. Given that the energy implications of not having and not using an appliance are basically the same, it was preferred to code not owning an appliance as zero usage per week (note that for some appliances, this might introduce some inaccuracy because of standby energy use; however, this was considered the lesser issue). For hours of daily TV usage, the usage across all TVs was summed; resulting in estimates up to 32 hours. Whilst separate estimates for TV use on weekdays, Saturdays, and Sundays, and tumbler dryer use in winter and summer were available, the respective measures correlated so highly (all correlation coefficients $r > .801$) that it made sense to calculate instead the average values, i.e. TV use on any day, and tumble dryer use in any week. To calculate the hours of lighting in the three rooms in winter, the hours across all light sets per room were summed.

Table 1
Overview of building variables and their frequencies. (bold = reference category).

Variable (abbreviation)	Categories (*N*)
Floor area (FloorArea)	n/a (continuous: $M = 93.82$ m^2, $SD = 43.69$ m^2)
Dwelling type (DwType)	Converted & purpose built flat (109), **detached (214)**, end terrace (106), mid-terrace (152), semi-detached (264)
Number of storeys (NoStorey)	n/a (continuous: $M = 2.01$; $SD = 0.71$)
Government Office Region (GOR)	East (88), East Midlands (58), London (89), North East (64), North-West (152), South East (113), South-West (84), West Midlands (77), Yorkshire and the Humber (120)
Dwelling age (DwAge)	**Pre 1919 (113),** 1919–44 (151), 1945–64 (198), 1965–80 (193), 1981–90 (60), post 1990 (130)
Wall type (WallType)	9-in. solid wall (118), cavity uninsulated (253), **cavity with insulation (421),** other (53)
Double glazing (DblgGlaz)	**Entire house (679),** more than half (95), less than half (34), no double glazing (37)
Attic (Attic)	Yes (93), **no (752)**
Conservatory (Conservatory)	Yes (176), **no (669)**
SAP rating (SAP)	**B&C (114),** D (492), E (209), F&G (30)

Table 2
Socio-demographic variables and their frequencies. (bold = reference category).

Variable (abbreviation)	Categories (*N*)
Number of occupants (HHSize)	**1 (197)**, 2(330), 3 (145), 4 (107), 5 or more (66)
Age of youngest dependent children (DepChild)	**No dependent children (562)**, 0–4 years (118), 5–10 years (76), 11–15 (61), older than 16 (28)
AHC (After-Housing-Costs) equivalized income quintiles (Income)	**1st quintile – lowest (121),** 2nd quintile (180), 3rd quintile (180), 4th quintile (104), 5th quintile-highest (184)
Tenure (Tenure)	Local authority (100), **owner occupied (558),** private rented (73), Registered Social Landlord RSL (114)
Sex of HRP (SexHRP)	**Female (337),** male (508)
Age of HRP (AgeHRP)	16–29 yrs (38), 30–44 (213), 45–64 (360), **65 or over (234)**
Employment status of household (EmployHH)	**1 or more work full time (432),** 1 or more work part time (74), none working and none retired (77), none working, one or more retired (262)
Someone in household sick or disabled? (sick/disabled)*	**No (553),** yes (292)
Someone in household over 75 years?	**No (751),** yes (94)
Length residency (LengthRes)	**2 yrs or less (136),** 3–4yrs (101), 5–9 years (175), 10–19 (189), 20–29 (119), 30 + years (125)

Table 3
Appliances and lighting – summary statistics of continuous variables.

Variable (abbreviation)	Mean (SD)	Median (1st–3rd quartile)
Set of lights kitchen	1.98 (1.13)	2.00 (1.00; 3.00)
Set of lights living room	2.78 (1.32)	3.00 (2.00; 4.00)
Set of lights main bedroom	2.16 (0.97)	2.00 (1.00; 3.00)
Hours lights on kitchen winter	5.60 (4.00)	4.00 (2.00; 7.00)
Hours lights on living room winter	7.15 (4.32)	6.50 (4.50; 9.00)
Hours lights on bedroom winter	1.86 (2.34)	1.00 (0.00; 3.00)
Number bulbs main kitchen light	2.93 (2.18)	3.00 (1.00; 4.00)
Number bulbs main living room light	2.88 (2.07)	3.00 (1.00; 4.00)
Number bulbs main bedroom light	1.59 (1.40)	1.00 (1.00; 1.00)
Number of TVs	2.43 (1.33)	2.00 (1.00; 3.00)
Hours TV watched any day	8.30 (5.00)	7.33 (4.67; 10.67)
Laundry loads per week	5.07 (4.30)	4.00 (2.00; 7.00)
Dishwashing loads per week	1.93 (2.87)	0.00 (0.00; 4.00)
Tumble dryer use per week	1.40 (2.66)	0.00 (0.00; 2.00)

Table 4
Categorical appliance and lighting variables (bold = reference category).

Variable (abbreviation)	Categories (*N*)
Any clothes dryer	**Yes (529),** no (316)
Any dishwasher	**Yes (370),** no (475)
Separate freezer	**Yes (415),** no (430)
Any microwave	**Yes (706),** no (139)
Supplementary electric heating	**No (766),** yes (79)
Any lighting on overnight?	**No (698),** yes(147)
Any low energy bulbs kitchen?	**No (580),** yes (265)
Any low energy bulbs living room?	**No (269),** yes (566)
Any low energy bulbs bedroom?	**No(284),** yes(561)
Weekly usage of electric oven	**Never (307),** 1 or 2 times (78), 3 or 4 times (145), 5 or 6 times (137), 7 or 8 times a week (147), 9 or more (31)
Weekly usage of electric grill	**Never (241),** 1 or 2 times (323), 3 or 4 times (153), 5 or 6 times (46), 7 or more (82)
Weekly usage of electric hob	**Never (595),** 1 to 4 times (41), 5 or 6 times (42), 7 or 8 times a week (115), 9 or more (52)

Table 5
Other occupant variables – summary statistics of continuous variables.

	Variable (abbreviation)	*M* (*SD*)
Answer scale	Do you agree that...	
1. Agree strongly	The Government is taking sufficient action to tackle climate change? (*Government*)	3.21 (1.04)
2. Tend to agree	It would embarrass me if my friends thought my lifestyle was purposefully environmentally friendly? (*Embarrass*)	3.07 (1.07)
3. Neither agree nor disagree	Being green is an alternative lifestyle, it's not for the majority? (*BeingGreen*)	3.06 (1.22)
4. Tend to disagree	I find it hard to change my habits to be more environmentally-friendly? (*Habit*)	3.31 (1.19)
5. Disagree strongly	It's not worth me doing things to help the environment if others don't do the same? (*NotWorth*)	3.63 (1.27)
Answer scale	How often, if at all, do you personally...	
1. Always	Switch off lights when you are not in the room? (*LightsOff*)	1.64 (0.98)
2. Very often	Boil the kettle with more water than you are going to use? (*BoilKettle*)	3.73 (1.30)
3. Quite often	Leave your TV or PC on standby for long periods of time? (*TVStandby*)	3.54 (1.63)
4. Occasionally	Wash clothes at 30 degrees or lower? (*Wash30*)	3.34 (1.58)
5. Never		

Note that all but ten households had an electric washing machine; hence, this variable was not included in the analysis.

Note that for the cooking variables 'never' included those households having the respective appliance powered by some other fuel (i.e. gas), not having or not using the appliance. The electricity consequences of those three cases are the same.

Table 5 shows those variables coded as 'other occupant' variables, measured on a Likert-scale.

Note that individual items are used as predictors instead of combining them into scales (e.g. construction of a "pro-environmental behaviour" scale). This was done as factor analysis and reliability analysis did not provide evidence for scales underlying the items. All items except for switching off lights were approximately normally distributed; that item was skewed to the right (median: 1.00; interquartile range 1.00–2.00).

The correlations between items were generally low, e.g. the mean correlation coefficient between the four items asking about energy-saving actions in the household was $r = .11$, ranging from $r = .004$ to $r = .215$. The items 'LightsOff' and 'Wash30' were reverse-coded for the correlation analysis (but not for subsequent regression analysis) so that positive correlations would be expected between all items.

One item was used as a categorical predictor, asking participants to indicate "Which of these statements best reflects how you currently feel?". The response options and number of participants who chose each option are summarized below. The part of the text printed in bold indicates how the respective item was later abbreviated in the results section.

- Climate change is caused by energy use and I'm beginning to think that I **should do something** ($N = 93$).
- Climate change is caused by energy use and I'm **doing a few small things** to help reduce my energy use and emissions ($N = 364$).

- Climate change is caused by energy use and I'm **doing quite a number** of things to help reduce my energy use and emissions ($N = 202$).
- Climate change is caused by energy use and I'm **doing lots of things** to help reduce my energy use and emissions ($N = 38$).
- I **don't believe** there are climate change problems caused by energy use and I'm **not willing or able to change my behaviour** ($N = 46$).
- Whether there are climate change issues or not, I **am not willing or able to change** my behaviour with regards to energy use ($N = 60$).
- Don't know ($N = 42$). – **don't know.**

2.3. Dependent variable: annualized combined energy consumption

The dependent variable used was the annualized electricity consumption in kilowatt-hours (kW h). The dependent variable was log-transformed (natural log) to achieve greater symmetry of the distribution, and of the residuals in the regression analysis. The mean log-transformed electricity consumption was $M = 8.18$ with a standard deviation of $SD = 0.59$. The geometric mean of the non-transformed energy consumption was $M = 3579$ kW h, and the arithmetic mean was $M = 4313$ kW h.[1]

2.4. Statistical analysis: ordinary least square and lasso regression

In a first step, a linear ordinary least squares (OLS) regression analysis was performed for each of the four classes of variables as presented above, i.e. 'building factors', 'socio-demographic', 'appliances and lighting', and 'other occupant variables'. Given the suspected issue of collinearity, the variance-inflation factors (VIF) were then inspected. VIF indicate how much the variance of an estimated regression coefficient increases if the explanatory variables are correlated. If uncorrelated, VIF = 1. There is no formal cut-off point for critical values of VIF; in this paper a value of 3.3 was used [33], this is a middle of the way value which is slightly higher than a conservative value of 2.5 (e.g. [34] but below other suggestions of above 5 or even 10 [35].

If VIFs greater than 3.3 were found in the OLS regression, then the Lasso regression (least absolute shrinkage and selection operator) was employed. Lasso regression is built on the linear model but uses a different procedure to calculate regression coefficients (see [35]; for an excellent description of this procedure, which was originally developed by [31]). Lasso is a penalised regression analysis promoting a sparser model. It uses a fitting procedure which sets some coefficients to zero, effectively aiding elimination of non-relevant variables. It aims to minimize the usual sum of squared errors, but constrained with a bound on the sum of the absolute values of the coefficients.

In order to choose the optimal tuning penalty parameter lambda λ (which penalises the sum of the absolute values of the regression coefficients), k-fold cross-validation was used, with 100 values for λ, and the data were randomly split into $k = 10$ groups. For each λ, the cross-validation error was calculated. Then the optimal value of λ was chosen which corresponds to the minimum cross-validation error (for details, see [35]. The "one-standard error" rule was applied; choosing as the final optimal value of λ that which gives the most regularized model (most sparse model) such that its error is within one standard error of the minimum error as estimated in cross-validation. After choosing the final value of λ, the model was re-run on all data.

Categorical variables were dummy-coded prior to analysis. Group-Lasso was used, which discards a categorical variable in total instead of individual categories within that variable to ease interpretation [36]; *R* package SGL).

After identifying which coefficients were set to zero using Lasso, then a new OLS was repeated omitting those variables.

After building all individual models (i.e. building, socio-demographics, appliances/ lighting, 'other occupant variables), models were then successively combined until resulting in a final model encompassing all predictors, tested and adjusted for multicollinearity.

3. Results

In the following section, the results of first the individual regression models (Sections 3.1–3.4) and the combined models (Section 3.5) are reported. Note that for all models the residuals were inspected. Inspection of the QQ plots of the residuals show that the residuals are nearly normal except for some outliers at both ends and that they are linear over a wide range of values. Furthermore, the residuals versus fitted values indicate that the residuals are nearly uncorrelated to the fitted values. For brevity, not all residual plots are presented. Appendix A shows an example of the residual plots for the final model (Section 3.5).

3.1. Building variables regression model

Building variables explained $R^2 = 16.7\%$ of the variability in log-transformed energy consumption, adjusted $R^2 = 13.6\%$, $F(30, 814) = 5.43$, $p < .001$. However, inspection of the VIF values showed multiple values above the cut-off of 3.3, making it necessary to run a Lasso regression.

In the Lasso regression, the following variables were set to zero: GOR, wall type, double glazing, SAP. These variables were hence excluded and the OLS regression rerun on the remaining variables. This reduced model explained $R^2 = 14.7\%$ of the variability in electricity consumption, adjusted $R^2 = 13.5\%$, $F(12, 832) = 11.97$, $p < .001$.

Table 6 shows the coefficients of the Lasso regression (β_L), and then of the reduced OLS (unstandardized coefficients B_{OLS}, standard error of unstandardized coefficients SE_{OLS}, standardized coefficients β_{OLS}). The stars indicate significance at the .001 (∗∗∗), .01 (∗∗), and .05 (∗) level in this and all subsequent tables. Unstandardized regression coefficients are in in the original measurement units, e.g. for floor area, it tells us how much energy consumption increases when floor area increases by one m^2. Hence, unstandardized coefficients are highly dependent on the scale of the independent variable. To allow comparison of impact of predictors measured in different units, standardized regression coefficients are stated. The standardized coefficient β refers to the number of standard deviation changes that are to be expected in the outcome variable for a one standard deviation change in the predictor variable. The significance level does not change.

Only two variables are significant: A larger dwelling size is associated with higher electricity consumption, and flats were associated with using less electricity than detached dwellings.

3.2. Socio-demographic regression model

The socio-demographic model explained $R^2 = 22.2\%$ (adjusted $R^2 = 19.4\%$) of the variability in residential electricity consumption, $F(29, 815) = 8.00$, $p < .001$. However, four variables showed VIF values above the chosen threshold criterion. Hence, Lasso regression was performed on the data. Five variables were set to zero: Pres-

[1] In regression models where the dependent variable has been log-transformed and the predictors have not, the format for interpretation is that dependent variable changes by 100 ∗ (coefficient) percent on average for a one unit increase in the independent variable while all other variables in the model are held constant (http://www.ats.ucla.edu/stat/sas/faq/sas_interpret_log.htm).

Table 6
Coefficients of the Lasso and OLS regression, building variables.

Predictor	β_L	B_{OLS}	SE_{OLS}	β_{OLS}
Floor area***	4.311	0.004	0.001	0.276
Dwtype (Ref = Detached): Flats*	−0.538	−0.205	0.080	−0.116
Dwtype: EndTerrace	0.077	0.011	0.074	0.006
Dwtype: MidTerrace	−0.058	−0.044	0.073	−0.029
Dwtype: Semi	0.135	−0.030	0.060	−0.023
Number Storeys	0.000	n/a	n/a	n/a
GOR (Ref = East): Midlands	0.000	n/a	n/a	n/a
GOR: London	0.000	n/a	n/a	n/a
GOR: North East	0.000	n/a	n/a	n/a
GOR: North-West	0.000	n/a	n/a	n/a
GOR: South East	0.000	n/a	n/a	n/a
GOR: South West	0.000	n/a	n/a	n/a
GOR: WestMidlands	0.000	n/a	n/a	n/a
GOR: Yorkshire & Humber	0.000	n/a	n/a	n/a
Dwage (Ref = pre1919): 1919–44	0.031	0.072	0.072	0.047
Dwage: 1945–64	0.087	0.136	0.070	0.097
Dwage: 1965–80	−0.093	−0.030	0.071	−0.021
Dwage: 1981–90	−0.084	−0.112	0.094	−0.049
Dwage: post1990	0.067	0.112	0.077	0.068
Wall (Ref = Cav. ins): Solid	0.000	n/a	n/a	n/a
Wall: Cavity uninsulated	0.000	n/a	n/a	n/a
Wall: Other	0.000	n/a	n/a	n/a
Double glazing (Ref = all): More than half	0.000	n/a	n/a	n/a
Double glazing: Less than half	0.000	n/a	n/a	n/a
Double glazing: None	0.000	n/a	n/a	n/a
Attic (1 = yes)	0.047	0.086	0.065	0.045
Conservatory (1 = yes)	0.293	0.097	0.050	0.066
SAP: D (Ref = B&C)	0.000	n/a	n/a	n/a
SAP: E	0.000	n/a	n/a	n/a
SAP: *F*&G	0.000	n/a	n/a	n/a
Intercept***	n/a	7.797	0.111	n/a

Table 7
Coefficients of the Lasso and OLS regression, socio-demographic variables.

Predictor	β_L	B_{OLS}	SE_{OLS}	β_{OLS}
HHsize 2 (Ref = HHsize 1)***	1.619	0.314	0.050	0.258
HHsize 3***	2.347	0.499	0.064	0.317
HHsize 4***	3.210	0.668	0.072	0.375
HHsize 5 or more***	3.232	0.798	0.085	0.362
DepChild(Ref = none): 0–4 years	0.000	n/a	n/a	n/a
DepChild: 5–10 years	0.000	n/a	n/a	n/a
DepChild: 11–15 years	0.000	n/a	n/a	n/a
DepChild: >16 years	0.000	n/a	n/a	n/a
Income	−0.102	−0.008	0.063	−0.006
Income2	−0.028	0.025	0.063	0.017
Income3	−0.026	0.020	0.064	0.014
Income4*	0.183	0.133	0.064	0.093
Tenure (Ref = Owner occ) Local authority***	0.000	n/a	n/a	n/a
Tenure: private landlord	0.000	n/a	n/a	n/a
Tenure: RSL	0.000	n/a	n/a	n/a
Gender HRP (1 = female)	0.523	0.049	0.039	0.041
AgeHRP (Ref: >65 yrs): 16–29 yrs	−0.293	−0.117	0.100	−0.041
AgeHRP: 30–44 yrs	0.266	0.001	0.065	0.001
AgeHRP: 45–64 yrs*	0.830	0.110	0.054	0.092
Employment (Ref = min 1 full time): at least 1 part time	0.000	n/a	n/a	n/a
Employment: none working, none retired	0.000	n/a	n/a	n/a
Employment: none working, at least 1 retired	0.000	n/a	n/a	n/a
Sick or disabled person (1 = yes)	0.000	n/a	n/a	n/a
Person over 75 yrs (1 = yes)	−0.312	−0.059	0.067	−0.031
Length residency (Ref ⩽ 2 yrs): 3–4yrs	0.000	n/a	n/a	n/a
Length residency: 5–9 yrs	0.000	n/a	n/a	n/a
Length residency: 10–19 yrs	0.000	n/a	n/a	n/a
Length residency: 20–29 yrs	0.000	n/a	n/a	n/a
Length residency: 30+ yrs	0.000	n/a	n/a	n/a
Intercept	n/a	7.726	0.072	n/a

Table 8
Coefficients of the OLS regression, lighting and appliances.

Variable (abbreviation)	B_{OLS}	SE_{OLS}	β_{OLS}
Number of lights kitchen	0.016	0.019	0.030
Number of lights	−0.011	0.017	−0.023
Number of lights main bedroom	0.005	0.021	0.009
Hours lights on kitchen winter*	0.011	0.006	0.094
Hours lights on living room winter	0.001	0.007	0.008
Hours lights on bedroom winter	0.004	0.008	0.016
Number bulbs main kitchen light*	0.026	0.009	0.095
Number bulbs main living room light*	0.024	0.009	0.085
Number bulbs main bedroom light	0.012	0.013	0.028
Any lighting on overnight?	0.006	0.048	0.004
Any low-energy bulbs kitchen?	0.006	0.041	0.005
Any low-energy bulbs living room?	−0.028	0.042	−0.023
Any low-energy bulbs bedroom?	−0.036	0.042	−0.029
Any dryer(REF = yes)**	−0.123	0.040	−0.101
Any dishwasher(REF = yes)	−0.051	0.059	−0.043
Separate freezer(REF = yes)**	−0.106	0.035	−0.090
Any microwave(REF = yes)	−0.062	0.047	−0.039
Any electric heating(REF = yes)	0.003	0.060	0.001
Number of TVs**	0.047	0.017	0.106
Hours TV watched any day**	0.012	0.004	0.105
Dishwashing loads per week **	0.030	0.010	0.145
Laundry loads per week***	0.019	0.005	0.137
Tumble dryer use per week	0.011	0.009	0.050
Weekly oven 1–2 times	−0.091	0.068	−0.044
Weekly oven 3–4 times	−0.092	0.057	−0.059
Weekly oven 5–6 times	−0.077	0.057	−0.048
Weekly oven 7–8 times	−0.024	0.054	−0.015
Weekly oven 9 and more	0.123	0.098	0.039
Weekly grill 1–2 times	−0.044	0.043	−0.036
Weekly grill 3–4 times	−0.008	0.052	−0.005
Weekly grill 5–6 times	0.007	0.081	0.002
Weekly grill 7 and more	−0.051	0.066	−0.025
Weekly hob 1–4 times	0.101	0.085	0.036
Weekly hob 5–6 times	0.085	0.083	0.031
Weekly hob 7–8 times	0.052	0.055	0.030
Weekly hob 9 and more	0.094	0.076	0.038
Intercept	7.752	0.131	n/a

ence of dependent children, Tenure, Employment status of household, Presence of sick/disabled person, and length of residency. Omitting these variables and performing OLS using the remaining ones resulted in an $R^2 = 21.2\%$, adjusted $R^2 = 20.0\%$, $F(13,\ 831) = 17.19$, $p < .001$.

Three variables exercised a significant effect. A larger household size was associated with increased energy consumption, being in the highest as opposed to the lowest income category was associated with higher energy use; as was being in the age category 45–64 as opposed to greater than 65 years (see Table 7).

3.3. Lighting/appliances regression model

The third regression model consisted of variables related to usage and ownership of lighting and appliances. The overall model explained $R^2 = 34.2\%$ of the variability in electricity consumption; adjusted $R^2 = 31.2\%$, $F(36, 808) = 11.65$, $p < .001$. All VIF were smaller than 2.91; hence, Lasso regression was not necessary; Table 8 shows the coefficients of the OLS regression.

Two lighting related variables were significant: longer lighting hours in the kitchen in winter had a positive effect on electricity consumption as did a larger number of bulbs in the main living room light. In terms of appliance ownership possession of a separate freezer and of a tumble dryer were associated with greater electricity use, as was an increasing number of TVs. In terms of appliance usage, watching more hours of TV per day, doing more dishwashing loads per week, and more laundry loads per week were all associated with greater electricity consumption. The cooking variables had no significant effect.

Table 9
Coefficients of the OLS regression, other occupant variables.

Predictor	B_{OLS}	SE_{OLS}	β_{OLS}
Government	0.014	0.020	0.025
Embarrass	0.008	0.020	0.015
BeingGreen	0.017	0.017	0.035
Habit	−0.012	0.019	−0.025
NotWorth	0.017	0.017	0.036
LightsOff	0.018	0.021	0.029
BoilKettle*	−0.042	0.016	−0.092
TVStandby**	−0.039	0.013	−0.108
Wash30	−0.009	0.013	−0.024
Believe in CC & should do sth (Ref = believe & do) lots)	0.224	0.118	0.118
Believe in CC & doing small things*	0.208	0.103	0.174
Believe in CC & quite a number*	0.234	0.104	0.168
Don't know	0.250	0.133	0.092
Don't believe in CC & don't want to change*	0.299	0.136	0.115
Don't know about CC & don't want to change	0.208	0.126	0.090
Intercept***	8.123	0.186	n/a

Overall, appliance use explained much more of the variability in residential electricity consumption than either socio-demographic or dwellings characteristics. However, it needs to be noted, that the variables for appliance usage are likely related to household size, i.e. a larger household might have more TVs and might likely have more hours of TV watching. Section 3.5 shows the impact of appliance variables when controlling for socio-demographics.

3.4. Other occupant variables regression model

The final individual regression looked at the impact of attitudes and self-reported environmentally significant behaviours on electricity consumption. The OLS regression explained $R^2 = 4.2\%$ of the variability in electricity consumption, adjusted $R^2 = 2.5\%$, $F(15, 829) = 2.45$, $p = .001$, all VIF < 1.35, see Table 9 for the regression coefficients.

Two self-reported behavioural actions were negatively associated with electricity consumption; i.e. a higher value in the behaviour was associated with lower electricity consumption. Hence, those moving towards 'never' in leaving the TV on standby and in overfilling the kettle, had lower electricity consumption as indicated by the negative coefficient. The only other significant variable was the one related to belief in climate change and (not) taking action. Numerically, all categories were associated with greater electricity consumption than the reference category of believing in climate change and doing lots; and all p-values were smaller 0.10. However, given the chosen significance level of <.05, only three comparisons were significant.

Table 10
Lasso and OLS coefficients for the final combined regression model.

Predictor	β_L	B_{OLS}	SE_{OLS}	β_{OLS}
Floor area**	2.318	<0.000	<0.000	0.111
HHsize 2 (Ref = HHsize 1)**	0.147	0.047	0.047	0.100
HHsize 3**	0.648	0.059	0.059	0.136
HHsize 4***	0.716	0.072	0.072	0.139
HHsize 5 or more***	0.893	0.081	0.081	0.143
Person over 75 yrs (1 = yes)	−0.217	0.055	0.055	−0.039
Hours lights on kitchen winter*	1.335	0.004	0.004	0.075
Number bulbs main kitchen light	0.551	0.009	0.009	0.051
Number bulbs main living room light**	0.711	0.009	0.009	0.083
Any dryer(REF = yes)**	−1.652	0.039	0.039	−0.101
Any dishwasher(REF = yes)	−0.598	0.056	0.056	−0.034
Separate freezer(REF = yes)*	−0.391	0.034	0.034	−0.074
Number of TVs	1.736	0.016	0.016	0.062
Hours TV watched any day***	1.313	0.004	0.004	0.116
Dishwashing loads per week *	1.430	0.010	0.010	0.110
Laundry loads per week	1.622	0.006	0.006	0.061
Tumble dryer use per week	0.506	0.009	0.009	0.064
BoilKettle	−0.168	0.013	0.013	−0.044
Intercept	n/a	7.652	0.103	n/a

3.5. Combining the induvial regression models

In the next step, we combined the different models together to test for increments in explanatory power through adding additional variables. For the building and socio-demographic model, only the variables that had remained after the Lasso regression were included.

In a first step, we combined building variables and socio-demographic variables. This model, 'build_socio' explained $R^2 = 27.6\%$ of the variability in domestic energy consumption; adjusted $R^2 = 25.4\%$, $F(25, 819) = 12.47$, $p < .001$. This increase was significant, to the model with building variables only ($p < .001$), and in comparison to the socio-demographics model only ($p < .001$).

In the second step, the appliance variables were added to the 'build_socio' model. The 'build_socio_appliance' model explained $R^2 = 38.9\%$ of the variability in electricity consumption, adjusted $R^2 = 34.2\%$, $F(61, 783) = 8.18$, $p < .001$. This 10% increase in R^2 was highly significant, $p < .001$.

In the third step, the attitudinal variables were added to the ' build_socio_appliance' model. This final model, 'build_socio_appli ance_attitudes', explained $R^2 = 40.0\%$ of the variability, adjusted $R^2 = 34.0\%$, $F(76, 768) = 6.73$, $p < . 001$. Adding the attitudinal variables did not increase explanatory power significantly, $p = .528$.

Fig. 1(a) shows the adjusted R^2 of the individual models, and (b) of the combined models. Appliance-related variables explain by far

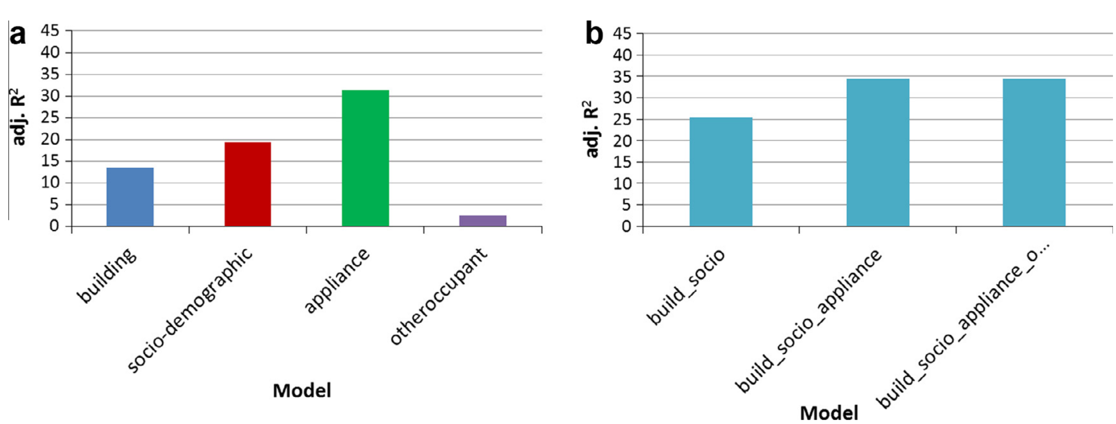

Fig. 1. Adjusted R^2 for the four individual models (a) and for the successively combined models (b).

the most of the variability in electricity consumption, on their own, and also when added to building and socio-demographic variables, they increase explanatory power by another 10%. Building variables play a lesser role in explaining electricity consumption when space and water heating is excluded.

The final model combining all variables showed VIF values above 3.3; hence, Lasso regression was run on the data to arrive at final coefficients.

Table 10 summarizes the coefficients for all variables that had remained after the Lasso regression. For brevity, those variables that were set to zero in the Lasso regression, are not shown. The model explained $R^2 = 36.0\%$ of the variability, adjusted $R^2 = 34.6\%$, $F(18, 826) = 25.84$, $p < .001$.

Only eight variables remained significant in this analysis. The only building variable that remained significant was dwelling size. As expected, household size was a strong predictor of electricity consumption. Two lighting related variables were significant, the hours of using the kitchen light in winter, and the number of bulbs in the main living room light. Ownership of a separate freezer and a dryer were associated with significantly higher electricity consumption. Hours of TV watched were highly significant; number of TVs owned was close to significance ($p = .081$). More dishwashing loads per week were also associated with higher electricity use.

4. Discussion

Our analysis focused solely on explaining electricity consumption without space heating and cooling. It is to the authors' knowledge the only study to be able to test such a large and varied number of predictors simultaneously in a nationally representative sample with more than 800 households. It is also one of the few studies to explicitly address multicollinearity and using an analysis technique to overcome this issue.

4.1. Summary and relation to previous research

To summarize, a total of 35% of the variability in electricity consumption was explained by the four classes of predictors (building variables, socio-demographics, lighting & appliance data, and 'other occupant variables'). The analysis showed that residential non-heating electricity consumption is to a large extent impacted on by appliance related variables and household size. Building variables played hardly any role, except for building size. Other studies show that when electricity-based heating is included, building variables play a much larger role [10–13], just as when looking at total energy consumption (e.g. [4,8,9]. Hence, depending on which part of energy consumption the interest lies, different variables need to be collected.

No relationship was found in the data with building age, insulation levels, SAP, and dwelling type which are factors that are presumably more important when looking at heating energy consumption because of their relationship to heat loss.

Regarding the impact of socio-demographic variables, overall, they explained electricity consumption better than building variables alone. For household size and composition, only household size was found to have a significant effect, but no effect was found of teenagers in the house, as previously reported by Brounen et al. [5] and Jones and Lomas [13]; in fact, this variable was set to zero in the Lasso regression analysis. Differences in sample and dependent variable (per capita consumption in Brounen et al.; combined gas and electricity in Jones and Lomas) might explain these differences regarding the effect of household composition; in addition, in our sample, the two variables household size and composition were highly related, hence creating an issue of collinearity which was resolved through Lasso regression setting one variable to zero. Regarding the question whether the effect of household size should be used a continuous predictor [10,12,20,21] or not [15,16,18,22], our data indicate the increase in electricity consumption with each additional household member becoming smaller as household size grows. The predicted values for electricity consumption using the coefficients from the final regression show that electricity consumption increases by an average of 1108 kW h when moving from a one to two-person household but by only 624 kW h when moving from a four-person to a 5-and more person household (note, for ease of understanding, predicted values were transformed back from the log scale by using the exponential).

Equivalized income only had a significant impact when looking at socio-demographic variables alone; once controlling for building and appliance data, the effect disappears. Note that non-equivalized income might show a different effect; however, equivalized income – i.e. adjusting a household's income for size and composition – allows to look at the incomes of all households on a comparable basis.

Appliance data on its own explained by far the largest share of variability in electricity consumption compared to the other individual models. Ownership of tumble dryers, separate freezers, and the frequency of use of appliances such as dishwashers played a significant role; whereas lighting related variables were of lesser importance. It might be that this is because lighting makes up a much smaller share of electricity consumption than consumer electronics or wet appliances [2]. Another potential reason is that only some relatively coarse variables for these parameters were used in this analysis (to avoid too small subsamples for very specific light bulbs), and that detailed lighting information might be harder to report than e.g. knowing whether one owns a tumble dryer.

Self-reported attitudes on climate change and pro-environmental behaviours, called 'other occupant variables' had an extremely small effect on electricity consumption when considered alone, and no effect when controlling for other variables. One reason might be that 'green lifestyles' are more commonly found in the higher income classes who might own more appliances, live in larger properties, i.e. having a larger energy consumption. Gilg et al. [37] found that income relates positively to self-reported green consumption; and higher recycling rates have been linked to higher income [38,39]. However, in our sample, there was no relationship between equivalized income and those 'other occupant' items. For example, treating equivalized income as a continuous variable (for ease of communication) showed a maximum correlation (absolute value) of $r = 0.11$ with the items on environmental opinions, and $r = 0.08$ with self-reported environmental actions. Hence, self-report might have simply been inaccurate and not reflect actual behaviour and lifestyles, potentially because of a social desirability bias [40], any impact might be too small to be picked up in electricity consumption, or finally, there might be other mediating variables (beyond income).

4.2. Implications of the study findings

The results imply that appliance ownership and usage is the most important variable explaining residential electricity use in properties not heated by electricity. Information about household size and dwelling size do add explanatory power of the models but an appliance only model already explains 31% (adjusted R^2) of the variability which only increased to 34% with demographic and building information. Hence, a detailed physical building characteristics survey would not necessarily be needed to understand electricity consumption.

In terms of future electricity consumption and the aim to reduce carbon emissions, it is important to ensure that further improvements of energy efficiency of appliances remain a priority

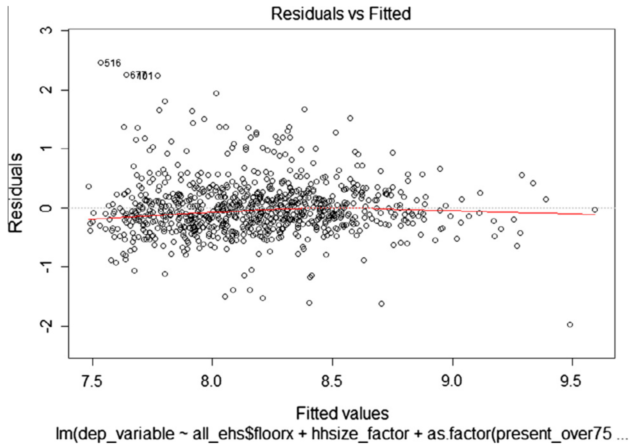

Fig. A1. Plot of fitted values against residuals for the regression model (as detailed in Section 3.5).

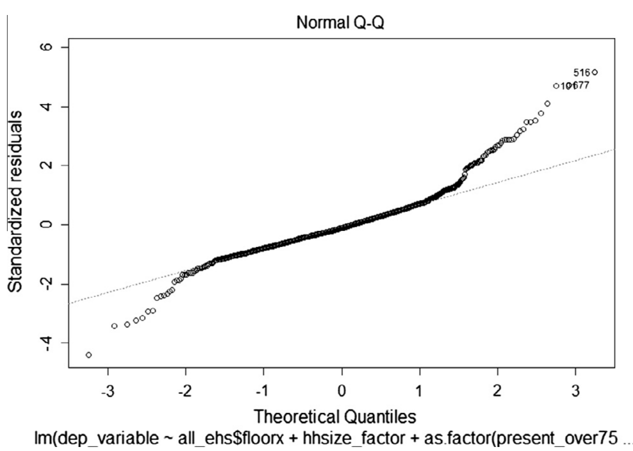

Fig. A2. Plot showing the normal Q–Q_ plot for the standardized residuals of the final regression model (as detailed in Section 3.5).

of product development and policy. In general, the energy efficiency of products has increased over the last decades [2], however, the potential energy savings are partly outweighed by owning more and larger appliances which has partly been accommodated by space per person increasing over time. One potential issue with a label like the EU energy efficiency label is that the energy efficiency label information might be more salient than the actual energy consumption which might lead people to buy a highly efficient but also high consuming appliance because of its size. Further research should aim at understanding better what drives choice of appliances (see e.g. [41]).

Hours of TV watched (controlling for household size) is associated with increased electricity consumption and has also been shown to be linked to childhood obesity [42–44], violent and aggressive behaviours in children [45,46], and greater risk of type 2 diabetes and cardiovascular disease in adults [47,48]. Hence, a joint campaign from different disciplines tackling TV use would have multiple beneficial outcomes.

Finally, the finding that owning a tumble dryer increases electricity consumption indicates that promotion and provision low-energy drying options is important. For example, when building new properties, in particular large blocks of flats, having a designated outside drying space, might encourage outdoor drying, as could the provision of an internal shared drying room (e.g. if dwellings are perceived to be too small for indoor air drying). Also, increasing the energy efficiency of dryers in particular ought to be on the forefront of product development.

In terms of methodological implications, the analysis has shown that multicollinearity can be an issue when studying determinants of electricity consumption. Multicollinearity leads to instable regression coefficients, meaning that an effect cannot be unambiguously ascribed to a variable. Hence, it is important that any analysis checks for multicollinearity and either chooses an appropriate analysis method, or care is exercised when interpreting results. Also, given that the impact of some seemingly important variables changes when controlling for other variables (such as equivalised income not being a significant predictor when other variables are controlled for), it is important to ensure that the effect of variables are not studied in isolation. Finally, the lack of impact of environmental concerns as surveyed and reported in the data analysed here questions the suitability of these types of items when trying to understand behaviour, in this case electricity consumption (but similarly for overall energy consumption, e.g. [4].

4.3. Limitations

The overall explanatory power of all variables together was limited; raising the question what other factors determine electricity consumption that were not measured in this study. Whilst the survey had included questions on other electrical appliances such as heated swimming pools, the numbers were generally too small to be analysed. Ownership and usage of personal computers was not assessed in the survey, and neither were the average annual number of weeks of vacation taken away from the house (e.g. Ndiayea and Gabriel [12]). However, it might be that other factors have a large impact on residential electricity consumption that are harder to assess quantitatively in a survey. In particular in sociological research, practice theory allows a much more detailed look at residential energy consumption (e.g. [49,50]. However, qualitative data is often limited to a small sample given how time consuming data collection and analysis tends to be which is then poorly representative, and in addition might not allow quantification of the effects of different variables. More recent approaches such as using smart-meter data to infer practices carried out in the house (e.g. [51]) might overcome these limitations and help to foster greater understanding of residential electricity use.

5. Conclusions

Using a large, nationally representative sample of 845 households, this paper showed that appliance ownership and usage are the most influential variables in understanding electricity consumption in gas-centrally heated buildings, together with household size. Hence, in order to reduce electricity consumption, energy-efficient appliances ought to become more and more widespread. Building variables played only a lesser role, as opposed to studies where total energy consumption including space heating is examined (e.g. [4]). Hence, depending on which part of energy consumption the research focus lies, different variables ought to be collected. Other occupant variables such as climate change concern and self-reported energy-relevant behaviours played hardly a role in understanding electricity consumption. Whilst this might give the impression, that trying to change attitudes towards the environment is futile, that conclusion would be far too preliminary. For example, it might well be that people with high environmental concern would be more likely to purchase energy-efficient appliances when being encourage to do so.

The study has important methodological implications, i.e. that checking for and addressing multicollinearity is crucial in performing regression analysis. In addition, how variables are coded, e.g. income and household size as categorical or continuous variables, deserves careful consideration given that results might differ depending on this decision.

Acknowledgements

Funding for this research was provided by the Research Councils UK (RCUK) Centre for Energy Epidemiology, grant reference EP/K011839/1.

Appendix A

(see Figs. A1 and A2).

References

[1] Palmer J, Cooper I. United Kingdom Energy Housing Factfile 2013. Retrieved from <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/345141/uk_housing_fact_file_2013.pdf> [accessed April, 1st, 2015].
[2] DECC. Energy Consumption in the UK (2015). Chapter 3: Domestic energy consumption in the UK between 1970 and 2014. <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/449134/ECUK_Chapter_3_-_Domestic_factsheet.pdf> [accessed 28-09-2015].
[3] DECC. The UK low carbon transition plan: national strategy for climate and energy. Department of Energy and Climate Change, The Stationary Office, London; 2009.
[4] Huebner GM, Hamilton I, Chalabi Z, Shipworth D, Oreszczyn T. Explaining domestic energy consumption – the comparative contribution of building factors, socio-demographics, behaviours and attitudes. Appl Energy 2015;159:589–600.
[5] Brounen D, Kok N, Quigley JM. Residential energy use and conservation: economics and demographics. Eur Econ Rev 2012;56(5):931–45.
[6] Aydinalp M, Ugursal V, Fung AS. Modeling of the appliance, lighting, and space-cooling energy consumptions in the residential sector using neural networks. Appl Energy 2002;71(2):87–110.
[7] Yin Y, Mizokami S, Aikawa K. Compact development and energy consumption: scenario analysis of urban structures based on behavior simulation. Appl Energy 2015;159:449–57.
[8] 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;41:1223–32.
[9] Steemers K, Young Yun G. Household energy consumption: a study of the role of occupants. Build Res Inform 2009;37(5–6):625–37.
[10] Bartusch C, Odlare M, Wallin F, Wester L. Exploring variance in residential electricity consumption: household features and building properties. Appl Energy 2012;92:637–43.
[11] Baker KJ, Rylatt RM. Improving the prediction of UK domestic energy-demand using annual consumption-data. Appl Energy 2008;85(6):475–82.
[12] Ndiaye D, Gabriel K. Principal component analysis of the electricity consumption in residential dwellings. Energy Build 2011;43(2–3):446–53.
[13] Jones RV, Lomas KJ. Determinants of high electrical energy demand in UK homes: socio-economic and dwelling characteristics. Energy Build 2015;101:24–34.
[14] Filippini M, Pachauri S. Elasticities of electricity demand in urban Indian households. Energy Policy 2004;32(3):429–36.
[15] Kavousian A, Rajagopal R, Fischer M. Determinants of residential electricity consumption: using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behaviour. Energy 2013;55:184–94.
[16] Wiesmann D, Azevedo IL, Ferrão P, Fernández JE. Residential electricity consumption in Portugal: findings from top–down and bottom–up models. Energy Policy 2011;39(5):2772–9.
[17] Wyatt P. A dwelling-level investigation into the physical and socio-economic drivers of domestic energy consumption in England. Energy Policy 2013;60:540–9.
[18] Yohanis YG, Mondol JD, Wright A, Norton B. Real-life energy use in the UK: how occupancy and dwelling characteristics affect domestic electricity use. Energy Build 2008;40(6):1053–9.
[19] Tiwari P. Architectural, demographic, and economic causes of electricity consumption in Bombay. J Policy Model 2000;22(1):81–98.
[20] Bedir M, Hasselaar E, Itard L. Determinants of electricity consumption in Dutch dwellings. Energy Build 2013;58:194–207.
[21] Cramer JC, Miller N, Craig P, Hackett BM. Social and engineering determinants and their equity implications in residential electricity use. Energy 1985;10(12):1283–91.
[22] Mansouri I, Newborough M, Probert D. Energy consumption in UK households: impact of domestic electrical appliances. Appl Energy 1996;54(3):211–85.
[23] McLoughlin F, Duffy A, Conlon M. Characterising domestic electricity consumption patterns by dwelling and occupant socio-economic variables: an Irish case study. Energy Build 2012;48:240–8.
[24] Abrahamse W, Steg L, Vlek C, Rothengatter JA. A review of intervention studies aimed at household energy conservation. J Environ Psychol 2005;25:273–91.
[25] Darby S. The effectiveness of feedback on energy consumption. A review for DEFRA of the literature on metering, billing and direct displays. Oxford: Environmental Change Institute, University of Oxford; 2006.
[26] Nilsson A, Bergstad CJ, Thuvander L, Andersson D, Andersson K, Meiling P. Effects of continuous feedback on households' electricity consumption: potentials and barriers. Appl Energy 2014;122:17–23.
[27] Vringer K, Blok T. Household energy requirement and value patterns. Energy Policy 2007;35:553–66.
[28] Abrahamse W, Steg L. How do socio-demographic and psychological factors relate to households' direct and indirect energy use and savings? J Econ Psychol 2009;30:711–20.
[29] Brandon G, Lewis A. Reducing household energy consumption: a qualitative and quantitative field study. J Environ Psychol 1999;19:75–85.
[30] Huebner GM, Cooper J, Jones K. Domestic energy consumption – what role do comfort, habit, and knowledge about the heating system play? Energy Build 2013;66:626–36.
[31] Tibshirani R. Regression shrinkage and selection via the lasso. J Roy Stat Soc B 1996;58(1):267–88.
[32] DECC. Energy follow-up survey 2011. Report 9: methodology; 2013. <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/274780/11_Methodology.pdf> [accessed 08.03.15].
[33] Roberts N, Thatcher J. Conceptualizing and testing formative constructs: tutorial and annotated example. Database Adv Inform Syst 2009;40(3):9–39.
[34] Chan SH, Chen JH, Li YH, Tsai LM. Gly1057Asp polymorphism of insulin receptor substrate-2 is associated with coronary artery disease in the Taiwanese population. J Biomed Sci 2012;19:100.
[35] James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning with applications in R. 1st ed. New York, Heidelberg, Dordrecht, London: Springer; 2013.
[36] Simon N, Friedman J, Hastie T, Tibshirani R. A sparse-group lasso. J Comput Graph Stat 2012;22(2):231–45.
[37] Gilg A, Barr S, Ford N. Green consumption or sustainable lifestyles? Identifying the sustainable consumer. Futures 2005;37:481–504.
[38] Schultz WP, Oskamp S, Mainieri T. Who recycles and when? A review of personal and situational factors. J Environ Psychol 1995;15:105–21.
[39] Smallbone T. How can domestic households become part of the solution to England's recycling problems? Business Strat Environ 2005;14:110–22.
[40] Fisher RJ. Social desirability bias and the validity of indirect questioning. J Consum Res 1993;20:303–15.
[41] Wijaya ME, Tezuka T. Measures for improving the adoption of higher efficiency appliances in Indonesian households: an analysis of lifetime use and decision-making in the purchase of electrical appliances. Appl Energy 2013;112:981–7.
[42] Dietz WH, Gortmaker SL. Do we fatten our children at the TV set? Television viewing and obesity in children and adolescents. Pediatrics 1985;75:807–12.
[43] Gortmaker SL, Must A, Sobol AM, Peterson K, Colditz GA, Dietz WH. Televison viewing as a cause of increasing obesity among children in the United States 1986–1990. Arch Pediatr Adolesc Med 1996;150:356–62.
[44] Robinson TN. Reducing children's television viewing to prevent obesity: a randomized controlled trial. JAMA 1999;282:1561–7.
[45] Dodge KA, Bates JF, Petitt CS. Mechanisms in the cycle of violence. Science 1990;250:1678–83.
[46] Windom CS. Does violence beget violence: a critical examination of the literature. Psychol Bull 1989;106:3–28.
[47] Grøntved A, Hu FB. Television viewing and risk of type 2 diabetes, cardiovascular disease, and all-cause mortality: a meta-analysis. J Am Med Assoc 2011;305(23):2448–55.
[48] Wijndaele K, Brage S, Besson H, et al. Television viewing and incident cardiovascular disease: prospective associations and mediation analysis in the EPIC norfolk study. PLoS One 2011;6(5):1–9. article e20058.
[49] Gram-Hanssen K. Standby consumption in households analyzed with a practice theory approach. J Ind Ecol 2010;14(1):150–65.
[50] Morley J. The significance of difference: understanding variation in household energy consumption. European council for an energy efficient economy. Summer Study, Belambra Presqu'ile de Giends, France.
[51] Torriti J, Hanna R, Anderson B, Yeboah G, Druckman A. Peak residential electricity demand and social practices: deriving flexibility and greenhouse gas intensities from time use and locational data. Indoor Built Environ 2015;24:891–912.