Daily life and demand: an analysis of intra-day variations in residential electricity consumption with time-use data

Aven Satre-Meloy · Marina Diakonova ·
Philipp Grünewald

Abstract Demand-side flexibility has been suggested as a tool for peak demand reduction and large-scale integration of low-carbon electricity sources. Deeper insight into the activities and energy services performed in households could help to understand the scope and limitations of demand-side flexibility. Measuring and Evaluating Time- and Energy-use Relationships (METER) is a 5-year, UK-based research project and the first study to collect activity data and electricity use in parallel at this scale. We present statistical analyses of these new data, including more than 6250 activities reported by 450 individuals in 173 households, and their relationship to electricity use patterns. We use a regularization technique to select influential variables in regression models of average electricity use over a day and of discretionary use across 4-h time periods to compare intra-day variations. We find that dwelling and appliance variables show the strongest associations to average electricity consumption and can explain 49% of the variance in mean daily usage. For models of 4-h average "de-minned" consumption, we find that activity variables are consistently influential, both in terms of coefficient magnitudes and contributions to increased model explanatory power. Activities relating to food preparation and eating, household chores, and recreation show the strongest associations. We conclude that occupant activity data can advance our understanding of the

temporal characteristics of electricity demand and inform approaches to shift or reduce it. We stress the importance of considering consumption as a function of time of day, and we use our findings to argue that a more nuanced understanding of this relationship can yield useful insights for residential demand flexibility.

Introduction

The residential sector is the largest end user of electricity in the UK, accounting for 45% of total consumption in 2017 (BEIS 2018). It is also responsible for up to 50% of national peak demand, during which time electricity provision is especially costly and carbon-intensive (Ofgem 2010). As evidenced by the share of generation from renewables increasing to over 29% in 2017, the UK's electricity system is increasingly low-carbon, yet much more ambition is needed to achieve national climate goals in the next decades (BEIS 2018).

The UK's 2017 Clean Growth Strategy sets forth a set of policies and proposals to further accelerate the deployment of low-carbon energy while maintaining increased economic growth (BEIS 2017). One of the proposals, the "Smart systems plan," aims to help consumers use energy more flexibly. Energy storage will support integration efforts as the UK pursues rapid deployment of renewable energy, but demand flexibility

A. Satre-Meloy (✉) · M. Diakonova · P. Grünewald
Environmental Change Institute, University of Oxford, South Parks Rd, Oxford OX1 3QY, UK
e-mail: aven.satremeloy@ouce.ox.ac.uk

can reduce the cost of integration and storage requirements for low-carbon energy systems (National Infrastructure Commission 2016).

Understanding how to make demand more flexible in the residential building sector requires a more detailed investigation of what electricity is used for in households (Grunewald and Diakonova 2018a; Powells et al. 2014). Time-use data is increasingly used for modeling electricity consumption, but these analyses often use simulated activity data and proxies for electricity consumption, such as household expenditure data, which requires assumptions about the link between time-use data and electricity consumption patterns. In some cases, these models are validated against real consumption data, though the validation sample sizes are in the 10–20 household range (Richardson et al. 2010; Widén et al. 2009; Widén and Wäckelgård 2010). Even as time-use data has been more recently incorporated into these models, however, there remains a notable gap in the literature of empirical studies on how occupant activities influence the temporal aspects of demand (Anderson 2016; Anderson and Torriti 2018; Grunewald and Diakonova 2018b).

In this paper, we aim to contribute to filling this gap by constructing regression models of original survey, time-use, and electricity consumption data for 173 UK households collected as part of an ongoing study. We test the extent to which socio-demographic, dwelling, appliance, and activity data can explain variations in intra-day household electricity use, from 5 p.m. until 9 p.m. on the following day. We then compare the effects of reported time-use activities in households throughout the day by dividing the day into six equal 4-h periods.

We aim to understand whether and how the inclusion of categorized activity data improves models of household electricity consumption. We investigate how different categories of activities are associated with electricity consumption during different times of day, and we consider how this knowledge can inform efforts to enhance demand-side flexibility and enable more and less costly integration of renewable energy sources.

The paper is organized as follows: In the "Literature review" section, we present a brief review of the flexibility and household electricity modeling literature, including both determinants of household electricity consumption and the use of time-use data in energy modeling. In the "Data and methodology" section, we present our methodology for collecting household electricity and time-use data in parallel. The "Results" section presents results of full-day and 4-h models of average electricity consumption. In the "Discussion" section, we discuss the implications of our findings, and in the "Conclusions" section, we conclude.

Literature review

Demand-side flexibility

Flexibility has been defined in the energy demand literature as the ability for consumers to change how, when, and where energy is used—a shift in focus from the magnitude of overall consumption to the timing of demand (Powells et al. 2014; Torriti et al. 2015). We conceive of it here simply as "a potential to change power at a certain rate (Watt/hour)" (Grunewald and Diakonova 2018a, p. 59). While much of the focus in demand-side energy research has been on reducing demand through energy efficiency, as the energy system becomes increasingly supplied by variable, low-carbon generation, flexibility becomes increasingly important for balancing supply and demand. The potential responsiveness of demand, or its ability to shift in time to match high generation from renewables or to flatten demand during peak periods, is essential for minimizing the costs of transitioning to a low-carbon power system (Strbac et al. 2012).

Demand-side response (DSR) refers to measures that provide flexibility to the energy system by shifting or reshaping energy loads. McKenna et al. (2017) conceive of DSR as a three-dimensional space including "technology change," "service expectation change," and "activity change." The technology change dimension refers to an automated or remotely controlled shift in energy use, for example, a smart fridge that delivers a response without affecting the energy service delivered. The authors note this has been the main form of DSR that has been tested in energy models given the relative ease of simulating such technology changes. "Service expectation change" refers to shifting occupants' expected level of energy service from an appliance or technology (i.e., the thermostat set-point), while an "activity change" requires a shift in either the timing or type of activity and relies on more active behavioral changes among energy users.

Grünewald and Diakonova (2018a) more broadly differentiate these dimensions of DSR as "appliance led" and "activity led." Among "activity led" DSR are

several different types of shift mechanisms: activity shifts, or changing the timing of an activity; substituting practice, which does not reorder activities in time but instead substitutes an energy-intensive activity for one that is not (e.g., having a cold rather than hot meal); substituting metabolic energy, such as mixing dough by hand rather than with an electric mixer; or changing the practitioner, which might entail going out for dinner rather than cooking at home.

Recent research has begun to investigate the DSR potential of these and similar types of measures, but still little is known about both the mechanisms by which electricity consumption patterns can be reshaped and the different constraints or motivating factors that can re-shape them. A review of major DSR trials in the UK found that residential customers are responsive to eco-nomic incentives to shift demand but that the size of the shift can vary significantly (DECC 2012). The report notes several areas where findings were inconclusive. These include the responsiveness of vulnerable and low-income customers, the impact of non-economic signals, and detailed information on the way consumers shift their usage in response to incentives.

Further empirical work on the flexibility of house-hold activities includes research that qualitatively assessed the likelihood of certain practices being per-formed during peak hours (4–8 p.m.) (Powells et al. 2014). This study found a very high or high likelihood of TV watching, cooking, computer/Internet use, and dish washing during these hours, with lesser frequencies of laundry, ironing, vacuuming, and bathing. More in-sight into the types of activities happening during peak periods and their contributions to electricity demand can help identify those that are best suited for DSR.

But how well "suited" an activity is for shifting is not the same as how "flexible" it might be. Torriti et al. (2015) stress this point by developing a "Flexibility Index" composed of five separate indices (synchroniza-tion, variation, non-shared activities, active home occu-pancy, and spatial mobility). These component indices are calculated for five time periods throughout the day for a sample of 153 respondents, and the findings show that morning peak times feature high levels of synchro-nization and shared activities but low variation in activi-ties. Evening peak times show less synchronization but higher spatial mobility and variation in activities.

Demand-side flexibility is complex and is difficult to measure empirically. It is likely that socio-demographic, dwelling, appliance ownership characteristics, and

activity patterns bear a strong relationship to both the types of DSR that can take place in homes as well as the capacity of households to participate in DSR programs. Several different approaches have been developed to better understand these relationships in order to inform more effective DSR. The following sections review these approaches in the context of research on electricity demand modeling and incorporating time-use data in these models.

Household electricity demand: models and determinants of use

The literature on quantitative approaches to model and understand key socio-technical determinants of electric-ity use in households is expansive. Previous studies employ both top-down and bottom-up approaches. Top-down approaches consider how national-level data, such as characteristics of the housing stock and macro-economic factors, influence electricity consumption, while bottom-up approaches use detailed household-level data on dwelling characteristics and occupant socio-demographics to identify the primary drivers of electricity consumption for a sample of households and then extrapolate these findings to the wider population (Swan and Ugursal 2009). Studies are also varied in how they measure household electricity use, in some cases considering average annual consumption values while in others studying daily or hourly consumption.

A recent comprehensive review of the socio-demo-graphic, dwelling, and appliance-related determinants of household electricity consumption undertaken by Jones et al. (2015) shows that no fewer than 62 factors have been found to affect household electricity consumption. Twenty of these are consistently found to have a signif-icant, positive correlation with electricity use ("consistently" is defined by the authors as showing a statistically significant relationship in more than three studies). The socio-demographic factors include number of occupants, presence of teenagers, household income, and disposable income. The impact of occupant age, education level, and tenure type is inconclusive in the studies reviewed. The dwelling factors that are influen-tial include dwelling age, dwelling type, number of rooms, total floor area, and ownership of electric space heating and cooling systems. Appliance factors include ownership of a desktop computer, television, electric oven, refrigerator, dishwasher, washing machine, and

tumble dryer, as well as the overall number of appliances owned.

Huebner et al. (2016) confirm many of these findings in one of the few representative studies of electricity consumption in the UK residential sector. Using data from a sample of 845 English households from 2011 to 2012, they find that models containing only appliance ownership and use factors explain 34% of the variance in non-heating annual electricity consumption, whereas models containing socio-demographic variables explain only 21%, and models containing dwelling and other occupant variables are poor in explaining electricity use. Their model combining these factors shows that dwelling floor area, number of occupants, wet appliance ownership, and hours of TV watched per day are statistically significant predictors of annual consumption.

As our paper focuses specifically on factors affecting intra-day variations in electricity use, a brief review of findings from studies with a similar focus is included here. The number of these studies is fewer than those included in the review discussed above because in much of the previous research, advanced metering data was not as available or accessible. With the growth of smart meter adoption, however, these data and studies investigating them are becoming increasingly common.

In general, this body of work suggests that appliance ownership and usage factors have stronger relationships with variations in consumption than socio-demographic and dwelling factors. In a study monitoring 27 UK residential buildings' electricity consumption at 5-min intervals for a period of 2 years, Firth et al. (2008) find that an observed increase in usage between years is primarily attributable to increases in the consumption of standby appliances, such as televisions and other consumer electronics, and "active" appliances, such as kettles, washing machines, electric showers, and lighting.

McLoughlin et al. (2012) study the effect of dwelling and occupant characteristics for a representative sample of 4200 Irish dwellings on several dependent variables, including total electricity consumption over a 6-month period, mean daily maximum demand over that period, electrical load factor (ratio of daily mean to daily maximum electrical demand), and timing of peak demand. They find that several dwelling and socio-demographic factors, such as number of bedrooms, type of dwelling, and head of household age, are significant in explaining load factor variations (i.e., those that are "peakier" or "flatter"). They also find that timing of peak demand is

mostly driven by variables measuring reported appliance ownership and use, especially dishwashers, electric power showers, televisions, and desktop computers. Notably, the models are less able to explain load factor and timing of peak demand than total consumption and mean daily maximum demand. Key explanatory factors for understanding variations in electricity consumption at different times of day may therefore be missing in the models.

Exploiting a similar smart meter data set from the Irish Commission for Energy Regulation (CER), Anderson et al. (2016) examine the links between half-hourly electricity consumption data and household characteristics for the purpose of assessing the feasibility of using high-resolution electricity data to infer household characteristics to supplement census-taking efforts. They construct multi-level regression models to analyze numerous "profile indicators" that describe the temporal characteristics of load profiles, such as base load, mean load, 97.5th percentile load (as a proxy for peak load), and several other temporal parameters. Their results show that several household characteristics, such as number of occupants, household income, and presence of children, are useful predictors of several of these "profile indicators," especially those measuring overall magnitude of consumption.

Kavousian et al. (2013) examine structural and behavioral determinants of daily maximum and minimum demand for a data set of 10-min interval electricity readings and an extensive household survey of 1628 US dwellings in California. The study was non-random and had a sample biased toward high-income and well-educated participants. They find that while daily minimum consumption is influenced most by weather, location, and physical characteristics of the building, such as size and type of home, daily maximum demand is influenced by high-consumption, intermittent appliances, such as electric water heaters or tumble dryers. The authors note the clear difference in their results between drivers of daily minimum and maximum demand, explaining that peak demand is more dependent on the activity patterns that lead to use of high-consumption appliances. Minimum demand is driven primarily by locational and physical characteristics of the dwelling.

Several recent papers have aimed to investigate socio-demographic and physical dwelling determinants of whole load profile shapes rather than average values of electricity consumption across varying time periods

(McLoughlin et al. 2015; Rhodes et al. 2014; Viegas et al. 2016). These studies take an alternative approach by first clustering load profiles into representative profile classes and then using logistic regression to determine factors influencing membership in different load profile classes. While this is a different approach than the one used in this paper and in the studies reviewed above, it achieves a similar aim. The findings from these studies indicate that variables such as working from home, time spent watching television, age, and ownership of dishwashers and washing machines are most important for explaining the shape of a household's daily electrical load profile.

The literature investigating highly resolved variations in residential electricity consumption is growing alongside the availability of data, but many of the results discussed here are dependent on the sample nature and size and not necessarily generalizable to wider populations. It is also important to consider the climatic and social contexts in which these studies are undertaken, as these likely influence their findings.

Incorporating time-use data in models of electricity consumption

Social scientists have long researched the dynamics of daily life through studies of how people spend their time. Longitudinal studies of thousands of households have investigated questions of historical changes in time use and how these relate to trends in technological and social change (eurostat 2009; Gershuny 2008; U.S. Bureau of Labor Statistics 2018). Research in energy demand modeling has acknowledged the need to take account of the timing of household activities in order to more accurately represent electricity demand (Torriti et al. 2015). In the last decade, occupant activity data from time-use studies has been increasingly incorporated into high-resolution models of residential electricity consumption (Torriti 2014). These studies involve varying approaches for incorporating time-use data. In some studies, probabilistic simulations are used to develop models of active occupancy, occupant activity sequences, and appliance usage that are predictive of electricity load curves (Ellegård and Palm 2011; Richardson et al. 2010; Widén et al. 2009). In others, changes in electricity consumption over decades are linked to changes in time-use and household expenditure data, often using decomposition analysis, which attributes changes in these factors to overall changes in

electricity use (De Lauretis et al. 2017; Jalas and Juntunen 2015; Sekar et al. 2018). In many of these studies, models and results are validated with actual consumption measurements from a small sample of households (McKenna and Thomson 2016; Widén et al. 2009).

While this literature shows varying results for the links between time use and electricity consumption, some common themes emerge. Mealtimes and related activities are consistently found to be high-consumption activities (De Lauretis et al. 2017; Druckman et al. 2012; Jalas and Juntunen 2015). Studies also find evidence that housework and personal time are energy-intensive activities (Palmer et al. 2013; Widén et al. 2009). Unsurprisingly, sleeping and resting are low-consumption activities (Druckman et al. 2012; Jalas and Juntunen 2015), whereas recreational activities can be either high or low energy depending on whether they involve televisions and computers or reading and socializing. In terms of the temporal shifts in the timing of UK peak demand over the past 40 years, Anderson and Torriti (2018) find that this is mostly due to changes in food-related activities, in part due to changes in personal care and housework shifts, and little to do with changes in media use.

Using the specific case of laundry practices in the UK, Anderson (2016) highlights the role of societal change in time shifts for energy-using activities. Increases in labor market participation by more females in society is likely partially responsible for shifting laundry energy demand into weekday mornings and evening peak times but also into Sunday mornings, which are less problematic for providing energy. The author concludes that more analyses of the changes to routine energy-using practices are necessary for assessing the value of various demand response strategies.

A review of time-use modeling of electricity demand undertaken by Torriti (2014) identifies five limitations in the literature. First, time-use data must be aggregated across large numbers of households to be statistically significant. Second, time-use data are often sampled throughout the year, whereas peak electricity events often happen on specific days, such as during temperature extremes or national events. Third, the low frequency of large time-use surveys means the data quickly become outdated. Fourth, the timing of occupancy and typical activities in developed countries is less variable than other factors that influence consumption, such as

weather or appliance design. Fifth, modeling multiple-person households using time-use data, especially for modeling occupancy patterns, is much more challenging than for single-person households, which was confirmed empirically by Grünewald and Diakonova (2018b). The fourth point, however, is contested. Tabbone et al. (2016) found that even within countries, cultural differences in time-use can affect load profiles.

To these limitations, McKenna et al. (2017) add several additional failings of time-use demand models. First, they note that time-use studies were not originally designed for the purpose of modeling energy use, and they thus do not differentiate between "energy-intensive" or "low-energy" alternatives of the same activity. Meal time is a good example, as this can vary from relatively low energy (cold meal preparation) to very energy intensive (oven and electric hob use). Second, time-use data do not account for overlapping or "bundled" activities, such as those that might occur during multi-tasking. Third, the typical 10-min reporting window for activities may not capture energy-relevant yet shorter duration activities, such as boiling the kettle.

This study addresses the limitations highlighted in these two reviews through a new approach to collecting time-use data. To the authors' knowledge, this is the first study to collect time-use data and electricity readings directly from households in parallel at this scale. This research contributes to an emerging focus within energy demand studies on not only the magnitude but also the timing of electricity consumption in households and how it is influenced both by numerous socio-demographic and dwelling factors that have been studied in previous research, as well as by the activities of occupants, which may contribute additional insights for determining the timing of electricity use in homes and the potential for shifting these activities to provide more flexibility to the energy system.

Data and methodology

Data collection

Electricity readings and activity records are collected from UK households as part of an ongoing study (Grünewald and Layberry 2015). Participation is voluntary, and participants are recruited online, via radio, and through campaigns at selected community events. An incentive for participation is the chance to win a year's worth of free electricity.

When registering for the study, participants complete a survey wherein they provide detailed socio-demographic information along with data on dwelling characteristics and appliance ownership. Participating households are sent a parcel prior to their chosen date, which includes an electricity recorder, activity recorder(s), and an instruction booklet. The study encourages all household members over the age of eight to participate.

Participants are instructed to attach the electricity recorder below the household's electricity meter. The electricity recorder collects readings every second for 28 h, from 5 p.m. on their chosen day until 9 p.m. the following day. This study length is chosen in order to capture two typical peak demand periods and because the electricity recorders have a battery life of around 28 h per charge. Fuel type used for heating and cooking appliances is collected in the household survey.

Activities are recorded with a dedicated app pre-installed on individual devices. The app presents six options per screen that guide participants through recording their activities, starting with the location and concluding with the number of people partaking in the activity and one's enjoyment of it. Asking about enjoyment was found to increase participant retention in previous time-use research (Gershuny and Sullivan 2017). Figure 1 shows an example activity entry sequence. In contrast to paper-based time-use diaries (eurostat 2009), the app can guide participants to select energy-relevant details, such as appliance use, allowing for detailed descriptions of activities. Instead of capturing the duration of activities, as was done in previous time-use studies, the app records activities reported instantaneously. Users can record multiple activities in sequence and are also given an option to record the "end" of activities. While users are encouraged to report activities when they actually occur, entries can be made retrospectively and into the future. One test of validity for reporting accuracy suggests around 80% of activity records for boiling the kettle are reported within 10 min of the activity itself (see Grünewald and Diakonova (2018b) for more details). More details about the functionality of the app is discussed in Grünewald et al. (2017), and a description of the data storage and handling procedures can be found in Grünewald and Diakonova (2019).

Study sample

The study sample consists of 173 households and 447 individuals who together reported 6265 activities. The process for determining inclusion in the subsequent analyses is as follows:

- Household that own solar PV are excluded due to concerns about the validity of electricity readings.
- Households that did not complete the full survey are excluded. The small occurrence of missing data (11% or 23 cases) did not warrant more complex handling, such as multiple imputation, so these cases are deleted listwise.
- Electricity readings below 20 W are excluded from daily and 4-h averages, as this signals a failure to attach the electricity recorder properly.
- Activities that were reported either before the electricity meter starts recording or after it stops are excluded. Similarly, activities that are reported in an hour where a valid electricity reading is missing are excluded.
- Only activities reported in the home are included, since we are interested in those activities that have a direct influence on household electricity consumption.

Socio-demographic, dwelling, and appliance variables

The household survey includes questions on occupant socio-demographics, physical dwelling characteristics, and appliance ownership. In previous studies, these have shown to have the strongest associations with residential electricity consumption in the UK (Huebner et al. 2016; Jones and Lomas 2016; Wyatt 2013). Table 1 presents frequencies and descriptive statistics for these variables for both the sample and the wider UK population (where available). Categorical variables are dummy coded prior to analysis, and reference categories are bolded in Table 1. Estimated monthly electric bill measures participants' estimate of their electricity expenditure. While we would expect this to have a positive association to average daily electricity consumption, we suspect that monthly estimates on expenditure might vary widely from day-to-day variations in consumption. We include it in our analyses to understand how well occupant estimates of electricity consumption do in fact predict actual average daily consumption values.

The study is voluntary, and the sample is not fully representative of the general UK population. Selection biases include an underrepresentation of renters, low-income groups, and the elderly. Gas boiler ownership is slightly underrepresented, suggesting a higher prevalence of all-electric heating in the sample. These biases are likely to influence both overall consumption as well as the timing of demand, which is why we caution against overgeneralizing from these results.

Table 1 also lists the distributions of season and day of week that households participate. A majority of households (57%) participate during the winter season (November–February). Data collection is focused on this season because it coincides with UK annual peak demand (DECC 2014). Households are recruited to start the study on a weekday, though given the study period spans 2 days, 22% of the sample starts on a Friday and completes the study on a Saturday. We do not create separate models for season or day of week in order to preserve the sample size across models, but we do include these as candidate variables during model selection, which is further explained below.
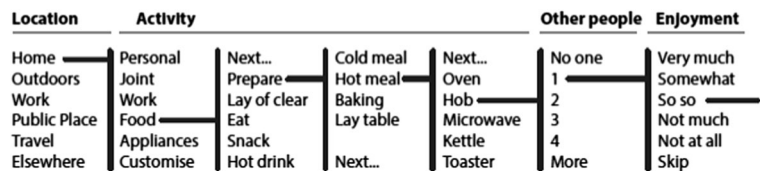
Activity variables

The app pre-codes activities, following the Harmonized European Time-Use Studies (HETUS) standards. For this analysis, activities are grouped into seven categories, which have become common in time-use research (Ellegård et al. 2010; Lader et al. 2006; Stankovic et al. 2016). Examples of activities that fall into each category are shown in Table 2.

In the following analyses, activities are aggregated across household members so that each household has a total count, or frequency, for each activity category. Table 2 presents descriptive statistics for each activity category for the sample.

Number and type of activities reported vary throughout the day, as shown in Fig. 2a. Most activities are reported in the early evening (5 p.m.–9 p.m.), followed by the morning period (5 a.m.–9 a.m.). For electricity research, this is helpful since early evenings also have the highest electricity use. Direct comparison of the early evening period, which is recorded on the first and second day, shows that the number of activities reported reduces by around 23%. Respondent fatigue is a possible explanation for some participants. The relative share of activities reported in each category

Fig. 1 Example of an activity entrance sequence on the activity recorder (Grünewald et al. 2017)

Location	Activity		Cold meal	Next...	No one	Very much
Home	Personal	Next...	Hot meal	Oven	1	Somewhat
Outdoors	Joint	Prepare	Baking	Hob	2	So so
Work	Work	Lay of clear	Lay table	Microwave	3	Not much
Public Place	Food	Eat		Kettle	4	Not at all
Travel	Appliances	Snack		Toaster	More	Skip
Elsewhere	Customise	Hot drink	Next...			

remains similar, however, so we do not expect this to bias the second 5–9 p.m. model.

Following Mckenna et al.'s (2017) point that the activity taxonomy developed for HETUS does not differentiate between "energy-intensive" and "low-energy" activities, we address this limitation by repeating the analysis described below including only activities we designate as "energy-intensive." Table 3 presents a list of these 26 activities and their descriptive statistics. These activities were reported a total of 2218 times by study participants. In Fig. 3, we show how the frequency of these activities varies throughout the day. While the total number of activities reported during each 4-h period remains proportional to the relative totals in Fig. 2a, the share of "food" and "recreation" activities is larger, and no "work" or "other care" activities are reported.

Dependent variable: average electricity usage
and average "de-minned" usage

We use several dependent variables in the following analyses. First, we model average daily electricity consumption in Watts (W) across the full 28-h study period. Next, we model average daily "de-minned" electricity consumption for both the full study period and for 4-h periods throughout the day. Electricity readings are averaged for the time period to which each model corresponds (i.e., the full-day models use the average and de-minned average from 5 p.m. on day 1 until 9 p.m. on day 2 while the other models use de-minned average electricity usage over each specified 4-h period). De-minning subtracts each household's minimum demand from its average usage in order to remove "baseload" electricity consumption (Jin et al. 2017). We use this technique because our aim is to characterize intra-day variations in electricity consumption and especially those variations that can be explained by activity data, which we expect to show stronger associations to de-minned electricity usage than normal averages. All dependent variables are log-transformed prior to regression analysis in order to improve linearity of regression residuals. Figure 2b shows boxplots for each dependent

variable prior to de-minning and transformation. Boxes indicate 25th, 50th, and 75th percentiles, while the whiskers indicate 1.5 times the interquartile range below or above the 25th or 75th percentiles, respectively. Dots represent outliers beyond these. Mean consumption during peak times is characteristically high compared to other day times. Some of the outliers in Fig. 2b are households with unusually high consumption throughout the day. As Fig. 2 suggests, number of activities reported and average electricity consumption are positively correlated during peak times, implying that activity reporting itself may be a useful proxy for understanding high usage.

The sample's mean daily electricity average of 565 W (SD = 340 W) is slightly higher than the "medium" estimates given by Ofgem's Typical Domestic Consumption Values (TDCV) for different classes of residential customers, which range from 350 to 490 W (Ofgem 2017). This may be because most households take part in winter months when electricity use is typically higher and because of the sample biases previously mentioned.

Statistical analysis: lasso and OLS regression

We use a variant on multiple linear regression to select final models of electricity consumption. This variant is the "least absolute shrinkage and selection operator" (lasso), developed by Tibshirani (1996). Lasso is one of the several prominent regularization techniques that have gained popularity for use in regression analyses where one or more of the following situations is present: the set of possible predictors is large and the analysis aims to identify those that most contribute to variations in the response variable; the data is "high-dimensional," meaning the number of predictors is larger than the number of observations; or the data suffers from a high degree of multicollinearity between predictors, which occurs when one predictor variable can be linearly predicted from the others in the model.

These situations are often present in electricity demand modeling. Failing to address them can lead to poor performance of the model and biased results due

Table 1 Selected socio-demographic, dwelling, appliance, and study participation date characteristics. National figures are based on estimates from various sources (CCC 2016; DECC 2013; ONS 2016, 2017, 2018a, b)

Description	Response	Sample frequency (%)	Population frequency (%)
Home type	Flat/apartment	20	21
	Detached	15	17
	Semi-detached	23	25
	Terraced	42	37
Tenure	Rent	20	32
	Own	80	64
No. of rooms, excluding kitchen and bathrooms	2 or fewer	6	4
	3–5	59	54
	6 or more	35	42
Household income	< £25,000	18	41
	< £35,000	8	16
	< £50,000	18	17
	> £50,000	56	27
No. of occupants	1–2	55	64
	3–4	41	30
	> 4	4	7
Age	Under 18	30	23
	19–50	52	44
	Over 50	18	35
Pet ownership	Cat	24	
	Dog	9	
	Fish	5	
Estimated monthly electric bill	< £400	40	
	< £700	41	
	> £700	19	
Bill affordability	Not affordable	15	
	Somewhat affordable	52	
	Affordable	33	
Electricity tariff	Renewable electricity	18	
	Economy 7 or 10	8	
	Other tariff	14	
	Do not know	14	
Large appliances	Washing machine	83	83
	Tumble dryer	31	50
	Washer dryer	16	14
	Underfloor heating	19	
	Gas boiler	83	85
	Heat pump	2	
	Electric stovetop	34	38
	Electric vehicle	0.3	0.4

Table 1 (continued)

Description	Response	Sample frequency (%)	Population frequency (%)
	Electricity display	9	
	Solar thermal	8	
Other appliances (count)	TV/computer screens	$M = 2.39$, $SD = 1.33$	
	Dehumidifiers	$M = 0.17$, $SD = 0.42$	
	Air conditioners	$M = 0.06$, $SD = 0.28$	
	Portable heaters	$M = 0.33$, $SD = 0.60$	
	Night storage heaters	$M = 0.12$, $SD = 0.67$	
	Power showers	$M = 0.32$, $SD = 0.65$	
Season participated	Winter (Nov–Feb)	46	
	Shoulder (Mar–Apr; Sept–Oct)	19	
	Summer (May–August)	35	
Day of week participated	Sunday–Monday	17	
	Monday–Tuesday	16	
	Tuesday–Wednesday	8	
	Wednesday–Thursday	36	
	Thursday–Friday	8	
	Friday–Saturday	16	

to unreliable model coefficients. Furthermore, conventional methods for variable selection, such as the stepwise methods, have been shown to cause various analytical issues, including overestimated R-squared values and regression coefficients and predicted values that are too narrow (Harrell 2001). Regularization techniques address these issues by constraining the magnitudes of coefficients and by trading a small increase in model bias for a larger reduction in variance (see Satre-Meloy (2019) for more details). While regularization techniques are the subject of a rich literature in statistics and machine learning and are well-suited to challenges in statistical modeling of electricity demand, their application in the energy literature thus far has been limited (Hsu 2015; Huebner et al. 2016).

Table 2 Activity categories, examples of their most frequently reported activities, and descriptive statistics

Time-use category	Example activities	Mean (SD)
Care for home	Arrange things, clear up, washing machine, vacuum cleaning, using tumble dryer	3 (2.7)
Care for others	Put child to bed, play with child, being with others, bathing children, with pet	3.2 (3.4)
Care for self	Sleep, shower, getting dressed, in bed, getting ready	7.7 (5.8)
Food	Hot drink, eat hot meal, cook on hob, wash dishes, eating	12.5 (8.2)
Recreation	Watching TV, reading, socializing, Internet, E-mail	7.3 (6.2)
Work	Work, main job, work second job, work with machinery, study	0.6 (1.1)

In this study, the data are neither "high-dimensional" ($n \gg p$) nor do the predictors suffer from a high degree of multicollinearity, which is determined by inspecting the variance inflation factors (VIFs) for the data.[1] For these reasons, the primary interest in applying regularization is to perform variable selection for a large predictor set to achieve a sparser model. While one of the other fundamental regularization techniques, ridge regression, performs well when dealing with "high-dimensional" data, it does not perform variable selection. The elastic net, a combination of ridge and lasso approaches, can address issues of multicollinearity but is somewhat more complex to implement (Zou and Hastie 2005). Our motivation in applying lasso regression is thus to attain sparsity and deliver statistical and computational gains (Tibshirani 2011).

Lasso applies a penalized linear regression model that shrinks the coefficients of some regression covariates while setting others exactly to zero, thus performing variable selection. While ordinary least squares (OLS) regression estimates coefficients by minimizing the residual sum of squares (RSS), lasso minimizes the RSS with an added penalty parameter based on $\sum_{j=1}^{p} |\beta_j|$ for some multiplier λ. Minimizing the RSS is thus given by the following equation:

$$\sum_{i=1}^{n} \left(y_i - \alpha - \sum_{j=1}^{p} \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \qquad (1)$$

The amount of penalty or shrinkage that is applied to the regression coefficients is controlled by the parameter λ. As λ is increased, a larger penalty is applied, and the estimates are progressively shrunk toward zero. In this way, lasso enables the selection of a model that does not overfit the data but still has low error. Both the predictors and the response are standardized to have mean zero and a standard deviation of one prior to running lasso.

Estimating the optimal value for λ is done using k-fold cross-validation, where models are constructed using a range of values for λ and each model's mean-squared error (MSE) after cross-validation is plotted. Comparing the MSE for each model at varying λ values then enables the selection of a parsimonious model with low MSE, a process known as hyperparameter tuning in machine learning disciplines. In cases where model sparsity is a primary aim, it is common to follow the "one standard error" rule, which says to select the simplest model that has an MSE within one standard error of the model with the minimum MSE (Friedman et al. 2010, p. 17). We follow this rule in the following analyses to improve model interpretability.

Lasso regression is applied to the full predictor set in order to construct a model for full-day average electricity consumption. The predictor variables selected at this stage are then included in an OLS regression, where both unstandardized and standardized coefficients are given. Unstandardized coefficients are measured in the original units of each independent variable. When the dependent variable has been log-transformed, it changes by 100× (unstandardized coefficient) percent on average for each one unit increase in the predictor variable, holding all other variables in the model constant. Standardized coefficients are instead measured in terms of standard deviation and thus can be compared in magnitude to determine which predictors have stronger or weaker associations with consumption given that predictors are measured in different units and on varying scales. Because our data contain numerous binary or dummy-coded predictors, we follow Gelman's (2008) suggestion of standardizing all non-binary predictors by

---

[1] VIFs signal whether regression coefficients are inflated due to correlation between predictor variables. If they are uncorrelated, VIF = 1. Traditionally, VIFs greater than 10 indicate high multicollinearity (Roberts and Thatcher 2009). In the models presented here, no VIFs are found to be greater than 10.
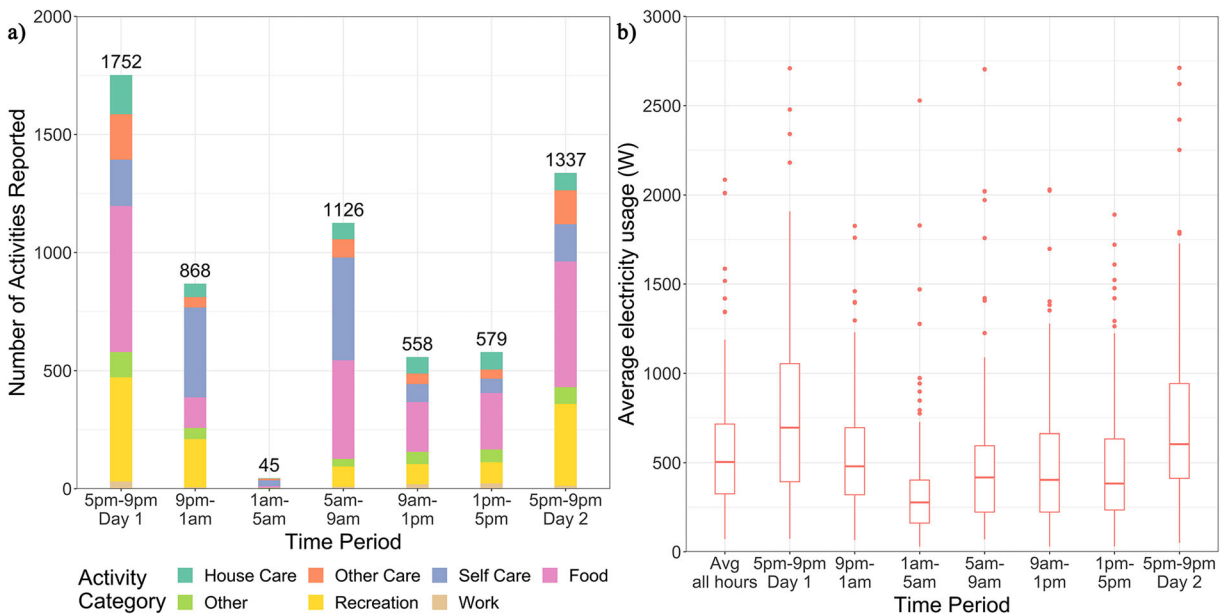
Fig. 2 a Frequency of activities reported by time period and time-use category ($N = 6265$). **b** Boxplots of the sample's average electricity consumption by time period ($N = 173$)

Table 3 Time-use codes, categories, and activity descriptions for "energy-intensive" activities

Time-use category	Time-use code	Activity description	Category mean (SD)
Care for home	3212	Vacuum cleaning	0.67 (1.02)
	3311	Washing machine	
	3320	Ironing	
Care for self	310	Wash and dress	1.73 (1.57)
	311	Shower	
	315	Hygiene/beauty	
	316	Wash	
	321	Bathing	
	8574	Shave	
Food	212	Eating hot meal	4.79 (3.70)
	214	Hot drink	
	3110	Prepare meal	
	3112	Prepare hot meal	
	3120	Baking	
	3124	Toaster	
	3133	Dish washer	
	3190	Prepare meal	
Recreation	3720	Internet	5.16 (4.63)
	3725	Online shopping	
	7170	E-mail	
	8001	Social media	
	8215	Watching TV	
	8221	Screen time	
	8552	Computer games	
	8553	TV	

Fig. 3 Frequency of "energy-intensive" activities reported by time period and time-use category ($N = 2218$)
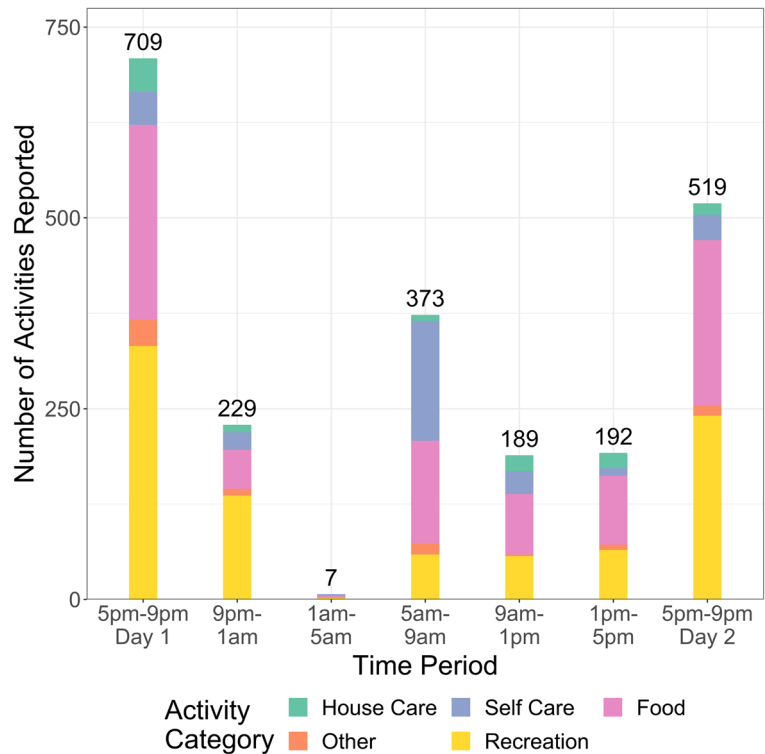
dividing by two standard deviations rather than one. This allows for numeric predictors' coefficients to be interpreted on the same scale as the binary inputs, which we leave unstandardized.

Next, we model daily average de-minned electricity consumption, again using lasso to select influential predictors. Finally, we use lasso to select separate models for each 4-h time period, starting with 5–9 p.m. on day 1 and ending with 5–9 p.m. on day 2. We choose a 4-h duration for each model to balance the aim of modeling typical patterns of daily life (i.e., peak demand hours, late evening, overnight, early morning, afternoon, etc.) while keeping the analysis from becoming overly complex, but our approach could be applied to consumption averages taken over shorter or longer time durations. We investigate the predictors selected in each model and compare their coefficients in order to identify which predictors are influential in explaining consumption patterns during different times of day. Next, using only activities we designate as "energy-intensive," we repeat this analysis to examine whether using a more energy-relevant categorization of activities improves model coefficients and/or explanatory power.

Finally, to investigate how much activity data contributes to explaining variations in de-minned average

electricity consumption (for both "all" as well as "energy-intensive" activities), we compare adjusted R-squared values for models with and without the activity variables included. *R Statistics* (R Development Core Team 2008) and the associated packages *dplyr* (Wickham et al. 2019) and *Glmnet* (Friedman et al. 2010) are used for data cleaning and regularized regression analyses. *Arm* (Gelman et al. 2018) is used to computed standardized coefficients, and *ggplot2* (Wickham 2016) is used for plots.

Results

The following section presents regression results of full-day and 4-h models of electricity consumption using household survey and activity data.

Lasso regression is performed on all survey and activity data to select influential variables in models of daily average electricity consumption and de-minned average electricity consumption. The penalty parameter λ is tuned using 10-fold cross-validation. Figure 4 shows the cross-validation procedure. The model with the minimum MSE (left dotted line) is found with a λ value of 0.03, includes 23 predictors, and has an MSE of

0.25. The model with an MSE within one standard error of the minimum (right dotted line) is found with a λ value of 0.08 and includes 13 predictors.

First, we include these selected 13 predictors in an OLS regression. Second, we re-run lasso on the full data, this time using de-minned average daily electricity consumption as the dependent variable. For the second model, lasso selects two new variables and sets two from the first model to zero, meaning the total number of selected predictors remains at 13. We again include these variables in a simple OLS regression. Table 4 presents regression results, including unstandardized coefficients (B) and 95% confidence intervals along with standardized coefficients (β) for both full-day models, with coefficients listed in order of standardized coefficient magnitude. For all regression models, we examine diagnostic plots of fitted values versus residuals and Q-Q plots. With few exceptions, these plots confirm normality and linearity of the residuals.

Model 1 in Table 4 explains $R^2 = 49\%$ (adjusted $R^2 = 0.44$) of the variance in average electricity consumption, $F(13, 159) = 10$, $p < 0.001$. The strongest predictors of increased daily consumption are mostly dwelling and appliance-related variables, especially EV ownership, number of power showers, living in a detached home, number of rooms, and number of TVs/computers. Socio-demographic variables that correlate to increased usage include cat ownership and number of occupants. The only activity category variable selected is recreation; number of recreation activities reported is associated with increased daily average consumption.

Ownership of a gas boiler and being on a renewable or "green" electricity tariff are the only variables in the model that predict decreases in average daily consumption. The former effect is likely observed because not owning a gas boiler indicates higher use of electricity for space and water heating. The latter effect may reflect a predisposition to conserve electricity among those who opt in to a renewable tariff.

The second model in Table 4, for which the dependent variable is de-minned average daily electricity consumption, explains $R^2 = 43\%$ (adjusted $R^2 = 0.38$) of the variance of the dependent variable, $F(13, 159) = 9.2$, $p < 0.001$. In addition to the variables in model 1, lasso selects being on an Economy 7 or 10 tariff, which typically charge lower prices for nighttime or "off-peak" electricity use, and number of night storage heaters. Both have positive coefficients but high uncertainty.

This model drops EV ownership and number of rooms as predictors.

In general, the differences between models 1 and 2 are slight, with model 2 showing slightly lower explanatory power than model 1. Notably, we do not find additional activity variables are selected in model 2, even though we might suspect activities to show a stronger association to de-minned rather than average daily consumption. We return to this finding in the "Discussion" section.

To further investigate how categories of time-use activities might explain patterns in household electricity consumption, we model de-minned average electricity use over 4-h time periods. Lasso regression is used to select models (which can have varying numbers of predictors), and we again include lasso-selected variables in an OLS regression for each 4-h period to investigate model coefficients. Table 5 shows variables and unstandardized coefficients with 95% confidence intervals. We exclude standardized coefficients from this table for simplicity but include these in a table in the Appendix.

Table 5 shows that for all times of day except 1 a.m.–5 a.m., activity category variables are frequently selected in these models and show strong associations to de-minned electricity consumption. Activities in the "care for home," "food," and "recreation" categories, in particular, are consistently selected as influential predictors across 4-h models.

Coefficients for activity variables show how different types of activities vary in their associations to electricity consumption at different times of day. Care for the home, which includes some energy-intensive household chores, has a stronger association with de-minned consumption in the early evening, mid-morning, and afternoon models and a weaker association in the late evening and early morning models. It is absent from the overnight model (1–5 a.m.) as well as from the day 2 5–9 p.m. model. Food-related activities follow patterns of typical mealtimes, with stronger coefficients in the evenings, early morning, and afternoon. The models do not show a consistent result in terms of the influence of different mealtimes on consumption, as the food coefficient for the morning mealtime is larger than in the day 1 5–9 p.m. model but smaller than in the day 2 5–9 p.m. model. This finding highlights that electricity consumption of food preparation and consumption can vary between days during the same mealtimes. Recreation activities show strong associations to electricity
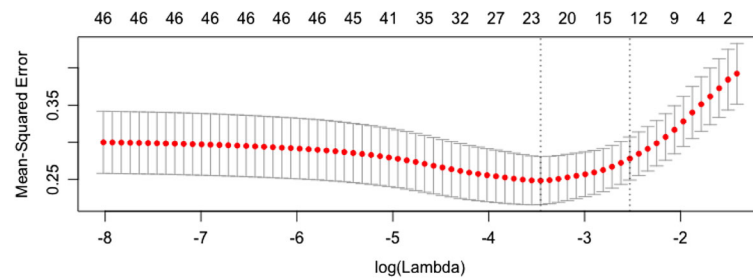
Fig. 4 Plot of cross-validation MSE for lasso regression of daily average electricity consumption. The horizontal bottom axis shows the logarithm of the tuning parameter λ, which increases in magnitude from left to right. The top horizontal axis shows the number of nonzero coefficients for each model run, and points and error bars show the mean and standard error of the cross-validation MSE, respectively. The left vertical dotted line indicates the λ value at which the minimum MSE is found, while the right vertical dotted line indicates the model with the fewest nonzero coefficients within one standard deviation of the minimum

consumption except overnight and during the early morning, but this variable is also absent from the day 2 5–9 p.m. model.

The only model in which no activity variables are selected is the overnight model, which also happens to be the model with the fewest selected predictors. Only EV ownership and being on an Economy 7 or 10 tariff are selected. Both of these can reasonably be expected to explain increases in nighttime electricity usage.

Regarding non-activity predictors that are selected in the 4-h models, many of these variables are the same as those selected in the models of 28-h electricity consumption. Some of these variables, such as living in a detached home, number of occupants, and estimated monthly electric bill appear to influence de-minned electricity consumption consistently throughout the day, as indicated by their selection in the majority of the models. Other variables are much more time-specific in their associations with de-minned electricity usage.

For instance, gas boiler ownership shows a large, negative association with electricity consumption in the late evening and early morning models. This finding may be intuitive, since space heating is typical in the late evening and early morning, and homes that heat with gas rather than electricity will likely have lower electricity consumption during these hours.

Because households select their own date for participation, we include season and day of week as candidate predictors for each model. Table 5 shows that there does not appear to be a notable day-of-week effect for our sample, but we do find a slight seasonal effect, as the 5–9 p.m. day 1 model includes the "participated in summer" predictor with a negative coefficient, and the 9 p.m.–1 a.m. model includes the "participated in spring or autumn" predictor with a positive coefficient. These

coefficients can be interpreted in terms of their comparison to the reference category, which is winter participation. Households that participate in summer show lower 5–9 p.m. de-minned electricity usage than households that participate in winter; similarly, households that participate in spring/autumn show slightly increased 9 p.m.–1 a.m. consumption compared with households participating in winter.

Regarding model explanatory power, we find that the model with the largest proportion of explained variance is the 5–9 p.m. day 1 model ($R^2 = 43\%$, adjusted $R^2 = 0.40$) and that the model with the lowest is the 5–9 p.m. day 2 model ($R^2 = 25\%$, adjusted $R^2 = 0.23$). The results for the two comparable 5–9 p.m. models also show inconsistencies in terms of variables selected, especially in the case of the activity variables. This finding challenges the notion of "average load profiles" (as commonly used for settlement, see Elexon (2013)) and suggests that activity-related factors that influence electricity consumption vary from day to day. This is a preliminary finding, however, and draws from a limited sample of days.

To address the limitation that these categories of activities do not differentiate between "energy-intensive" and "low-energy" activities, we repeat this analysis including only those activities that we designate as "energy-intensive." Table 6 presents results for these models. With the exception of several coefficients, we observe in most cases a marked increase in the care for home, food, and recreation activity variables' coefficients. This is especially true for "care for home" activities reported in the morning, mid-morning, and afternoon periods, suggesting energy-intensive housework is a strong predictor of de-minned usage during these times.

Table 4 Regression results for average electricity usage and average de-minned usage during full 28-h study period

	1) Daily average (W)		2) De-minned daily average (W)	
	B (5–95% CI)	β	B (5–95% CI)	β
EV	0.61** (0.18, 1.04)	0.61		
Tariff: Renewable	−0.30** (−0.50, −0.11)	−0.30	−0.30* (−0.54, −0.06)	−0.26
Gas boiler	−0.27* (−0.48, −0.06)	−0.27	−0.19 (−0.46, 0.09)	−0.16
Owns cat	0.25** (0.08, 0.43)	0.25	0.18 (−0.04, 0.39)	0.24
Tariff: Economy 7 or 10			0.43 (−0.03, 0.88)	0.24
Recreation activities	0.02** (0.01, 0.03)	0.24	0.03*** (0.01, 0.04)	0.23
No. of power showers	0.15* (0.03, 0.27)	0.19	0.13 (−0.02, 0.28)	0.18
Detached home	0.19 (−0.04, 0.43)	0.19	0.45** (0.18, 0.73)	0.34
No. of rooms	0.07* (0.005, 0.14)	0.18		
No. of TV/computers	0.06* (0.004, 0.12)	0.17	0.06 (−0.02, 0.13)	0.18
Under floor heating	0.16 (−0.03, 0.35)	0.16	0.16 (−0.07, 0.40)	0.20
Tumble dryer	0.10 (−0.07, 0.27)	0.10	0.10 (−0.11, 0.30)	0.08
No. of night storage heaters			0.07 (−0.11, 0.25)	0.01
Estimated monthly electric bill	0.03 (−0.02, 0.09)	0.10	0.02 (−0.05, 0.09)	0.13
No. of occupants	0.03 (−0.04, 0.10)	0.08	0.08 (−0.01, 0.16)	0.11
Constant	5.42*** (5.10, 5.74)	6.27	4.97*** (4.61, 5.33)	6.15
Observations	173		173	
R^2	0.49		0.43	
Adjusted R^2	0.44		0.38	
Residual std. error (df = 159)	0.47		0.57	

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$

In some instances, the energy-intensive activity models either drop or add activity-related variables for different times of day. While the "care for home" category when not filtering for energy-intensive activities is selected in the day 1 5–9 p.m. and the 9 p.m.–1 a.m. models, it is not selected when only energy-intensive activities are included. This may suggest that more energy-intensive activities related to household chores are not as common in the evenings. It also indicates that "care for home" as a category includes non-energy-intensive activities that are important for explaining electricity consumption during evenings.

Similar distinctions are found for food-related activities. The model including only energy-intensive food activities more clearly differentiates between evening and morning meals, as morning food activities show less associations with electricity consumption when excluding non-energy-intensive activities, likely because morning meals are often cold and do not involve appliance use. For recreation activities, the timing of when these are influential does not change between the two different categorizations, though the strength of associations clearly increases when only energy-intensive activities are included.

To test the extent to which categorized activity data improves model explanatory power, we compute the adjusted R^2 value for each model excluding the lasso-selected activity predictors for that model, and we do this for both categorizations tested.[2] We present these results in Table 7. We see that excluding activity data reduces model explanatory power by an average of nine percentage points across models (excluding the 1–5 a.m. model). We do not observe, however, a noticeable difference between the models including all activity data and those including only "energy-intensive" activities. These findings provide evidence of the potential improvements to household electricity modeling by including activity data, even when it is relatively coarsely categorized.

---

[2] Adjusted R^2 is used for the comparison in order to account for the varying number of predictors in each model.

Table 5 Regression results for de-minned average electricity use: 4-h models (5 p.m. day 1–9 p.m. day 2)

Predictor	5–9 p.m. (1)	5–9 p.m. (2)	9 p.m.–1 a.m.	1 a.m.–5 a.m.	5 a.m.–9 a.m.	9 a.m.–1 p.m.	1 p.m.–5 p.m.
Activity categories							
Care for home	0.17*** (0.09, 0.24)		0.16* (0.002, 0.32)		0.16 (−0.07, 0.39)	0.15 (−0.02, 0.33)	0.15 (−0.03, 0.32)
Care for others					0.15 (−0.04, 0.33)	0.35** (0.10, 0.60)	
Food	0.03 (−0.02, 0.07)	0.10*** (0.05, 0.15)	0.08* (0.01, 0.16)		0.06 (−0.01, 0.13)		0.08 (−0.02, 0.19)
Other activity						0.22* (0.02, 0.43)	
Recreation	0.06* (0.01, 0.12)		0.08* (0.01, 0.15)			0.21* (0.02, 0.39)	0.25* (0.04, 0.46)
Appliances							
EV				2.23*** (1.06, 3.40)			
Gas boiler			−0.37* (−0.65, −0.09)		−0.59** (−0.99, −0.20)		
Heat pump					0.76 (−0.20, 1.72)		
No. of night storage heaters							0.39** (0.15, 0.63)
No. of power showers			0.23** (0.06, 0.39)			0.27* (0.01, 0.53)	
No. of TVs/computers		0.07 (−0.04, 0.18)				0.17** (0.04, 0.30)	0.09 (−0.04, 0.22)
Tumble dryer					0.28 (−0.04, 0.60)		0.37* (0.01, 0.74)
Under floor heating			0.11 (−0.15, 0.37)		0.34 (−0.04, 0.72)		
Dwelling							
Detached home	0.62*** (0.29, 0.95)	0.56** (0.15, 0.97)	0.20 (−0.10, 0.50)		0.17 (−0.25, 0.59)		0.66** (0.19, 1.14)
No. of rooms in home		0.08 (−0.05, 0.20)				0.10 (−0.04, 0.23)	0.08 (−0.06, 0.22)
Terraced home	−0.16 (−0.41, 0.08)						
Socio-demographics							
Bill affordability	−0.18* (−0.35, −0.01)						
Estimated monthly electric bill		0.09 (−0.01, 0.19)	0.06 (−0.02, 0.13)		0.09 (−0.02, 0.19)	0.08 (−0.04, 0.21)	
No. of occupants	0.13* (0.03, 0.22)		0.06 (−0.03, 0.15)		0.09 (−0.06, 0.23)		0.13

Table 5 (continued)

Predictor	5–9 p.m. (1)	5–9 p.m. (2)	9 p.m.–1 a.m.	1 a.m.–5 a.m.	5 a.m.–9 a.m.	9 a.m.–1 p.m.	1 p.m.–5 p.m.
							(− 0.02, 0.27)
Owns cat		0.31 (− 0.01, 0.64)	0.36** (0.12, 0.59)		0.35* (0.02, 0.69)		
Participated in summer	− 0.43*** (− 0.67, − 0.18)						
Participated in spring/autumn			0.34** (0.09, 0.59)				
Tariff: Economy 7 or 10				1.73*** (1.07, 2.40)			
Tariff: Renewable	− 0.42** (− 0.72, − 0.11)		− 0.43** (− 0.69, − 0.17)				
Tenure (rent vs. own)					− 0.38* (− 0.73, − 0.02)		
Constant	5.82*** (5.38, 6.26)	4.55*** (4.00, 5.10)	5.04*** (4.68, 5.40)	3.92*** (3.75, 4.09)	4.53*** (3.98, 5.07)	3.53*** (2.91, 4.15)	3.42*** (2.79, 4.04)
Observations	171	170	172	172	172	170	170
R^2	0.43	0.25	0.43	0.28	0.33	0.29	0.39
Adjusted R^2	0.40	0.23	0.39	0.27	0.28	0.26	0.35

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$

Discussion

Summary and discussion of results

The results show that 13 lasso-selected predictors can explain 49% of full-day average household electricity usage. When the dependent variable is "de-minned" to remove baseload consumption, lasso selects a model that explains 43% of the variance in average de-minned electricity usage. These results are consistent with or in some cases improve on model explanatory power found in previous studies of household electricity usage (Huebner et al. 2016; Kavousian et al. 2013; McLoughlin et al. 2012). For the models of 4-h average de-minned electricity consumption, explanatory power is somewhat lower, ranging from 25 to 43% of variance explained. These models include varying numbers of lasso-selected predictors, but we find that removing activity variables from these models reduces their explanatory power considerably.

The results for both full-day models confirm previous findings that appliance ownership and occupant socio-demographics show the strongest associations to electricity consumption patterns (Huebner et al. 2016).

These results also, however, identify new variables that might help predict daily usage, such as electricity tariff choice or pet ownership. The results provide some evidence that these may be important for understanding electricity consumption patterns.

This study goes beyond previous research by investigating how the associations between appliance, dwelling, and socio-demographic variables and de-minned electricity consumption vary as a function of time of day. The EV ownership variable, for instance, is only selected in the nighttime model, which suggests that ownership of an EV increases a household's nighttime consumption more so than it increases consumption during other times. Similarly, the number of power showers and presence of underfloor heating are associated with consumption during late evening and morning hours, when showers are likely more frequent. Number of TVs/computers owned is a stronger predictor in the morning and evening.

The study also tests the extent to which different types of activities are associated with electricity consumption during different periods of the day. We find evidence that certain types of activities are more relevant for understanding electricity consumption patterns and

Table 6 Regression results for de-minned average electricity use: 4-h models (5 p.m. day 1–9 p.m. day 2) including only "energy-intensive" activities across categories

Predictor	5–9 p.m. (1)	5–9 p.m. (2)	9 p.m.–1 a.m.	1 a.m.–5 a.m.	5 a.m.–9 a.m.	9 a.m.–1 p.m.	1 p.m.–5 p.m.
Activity categories							
Care for home					1.05** (0.39, 1.71)	0.52* (0.13, 0.91)	0.70** (0.24, 1.15)
Care for self					0.11 (−0.05, 0.27)		
Food	0.02 (−0.06, 0.11)	0.18** (0.06, 0.30)	0.23** (0.06, 0.40)			0.14 (−0.07, 0.35)	0.20* (0.01, 0.39)
Recreation	0.10** (0.04, 0.16)		0.20*** (0.10, 0.29)			0.35** (0.12, 0.58)	0.30* (0.05, 0.55)
Appliances							
EV				2.23*** (1.06, 3.40)			
Gas boiler			−0.32* (−0.59, −0.05)		−0.65*** (−1.03, −0.27)		
No. of night storage heaters							0.42*** (0.18, 0.66)
No. of power showers			0.26** (0.10, 0.42)			0.22 (−0.05, 0.49)	
No. of TVs/computers						0.14* (0.01, 0.27)	0.08 (−0.05, 0.20)
Tumble dryer						0.34 (−0.03, 0.71)	0.38* (0.02, 0.75)
Under floor heating			0.12 (−0.13, 0.38)		0.40* (0.03, 0.77)	0.38 (−0.05, 0.81)	
Dwelling							
Detached home	0.68*** (0.34, 1.03)	0.66** (0.24, 1.07)	0.20 (−0.09, 0.50)		0.28 (−0.14, 0.69)	0.18 (−0.32, 0.69)	0.76** (0.28, 1.24)
No. of rooms in home		0.08 (−0.04, 0.21)				−0.01 (−0.16, 0.14)	0.07 (−0.07, 0.21)
Terraced home	−0.13 (−0.39, 0.13)						
Socio-demographics							
Estimated monthly electric bill		0.11* (0.01, 0.21)	0.04 (−0.03, 0.11)		0.10 (−0.01, 0.20)	0.04 (−0.08, 0.17)	
No. of occupants	0.16** (0.06, 0.26)		0.08 (−0.01, 0.17)		0.15* (0.01, 0.29)	0.12 (−0.04, 0.28)	0.15* (0.001, 0.30)
Owns cat	−0.43**		0.33** (0.10, 0.56)		0.33 (−0.01, 0.67)		

Table 6 (continued)

Predictor	5–9 p.m. (1)	5–9 p.m. (2)	9 p.m.–1 a.m.	1 a.m.–5 a.m.	5 a.m.–9 a.m.	9 a.m.–1 p.m.	1 p.m.–5 p.m.
Participated in summer	(− 0.69, − 0.17)						
Participated in spring/autumn			0.36** (0.12, 0.61)				
Participated on Fri–Sat						0.56* (0.11, 1.01)	
Tariff: Economy 7 or 10				1.73*** (1.07, 2.40)			
Tariff: Renewable	− 0.51** (− 0.81, − 0.20)		− 0.40** (− 0.65, − 0.14)				
Constant	5.56*** (5.24, 5.87)	4.77*** (4.24, 5.31)	4.99*** (4.63, 5.34)	3.92*** (3.75, 4.09)	4.52*** (4.03, 5.00)	3.62*** (2.97, 4.28)	3.48*** (2.86, 4.11)
Observations	171	170	172	172	172	170	170
R^2	0.35	0.20	0.45	0.28	0.30	0.33	0.37
Adjusted R^2	0.32	0.18	0.41	0.27	0.26	0.28	0.34

$*p < 0.05$; $**p < 0.01$; $***p < 0.001$

also that certain activities have stronger or weaker associations at different times of day.

In particular, our results show that housework, eating and meal preparation, and recreation or media use are strongly associated with de-minned electricity consumption, confirming similar findings from previous research (Anderson and Torriti 2018; De Lauretis et al. 2017; Jalas and Juntunen 2015; Palmer et al. 2013). The times at which these activities are stronger predictors of electricity consumption also provides insight into when activities and electricity use are more tightly coupled, such as during evening peak periods and later evenings for food and recreational activities, early mornings for personal care activities, and early and mid-mornings for household chores.

Our results also highlight the extent to which denoting activities as "energy-intensive" can strengthen these links. We find in general a sizeable increase in the strength of relationships between number of activities reported and average de-minned electricity consumption throughout the day when including only "energy-intensive" activities. We do note, however, that these results are somewhat mixed. For instance, the "care for home" variable is significant in the day 1 5–9 p.m. model when including all activity data but not when including only energy-intensive activities. This finding suggests other "care for home" activities are important for explaining electricity consumption during this time. While acknowledging that our designation of activities as "energy-intensive" is subjective, we take this finding as evidence that the links between activities and electricity consumption may be more complex than simply differentiating activities on the basis of their expected use of energy. Further discussion of this point can be found in Grünewald and Diakonova (2018b).

Two additional results merit further discussion. The first is our finding that we do not observe an increase in the number of activity variables selected between the daily average and daily de-minned average models. We do, however, observe frequent selection of activity variables in the 4-h de-minned usage models. We expect that this observation results from the duration of time over which we are de-minning electricity consumption. Over a 28-h period, de-minning does not appear to yield stronger associations between activities reported and average consumption. Over shorter 4-h periods, however, de-minning does appear to strengthen relationships between reported activities and consumption. This finding suggests that activity data can improve models of shorter duration more so than it can models of longer

Table 7 Model explanatory power with and without lasso-selected activity variables for both "all" activity data and "energy-intensive" activity data models

	5–9 p.m. (1)	5–9 p.m. (2)	9 p.m.–1 a.m.	1 a.m.– 5 a.m.	5 a.m.– 9 a.m.	9 a.m.–1 p.m.	1 p.m.–5 p.m.
"All" activity models							
Adjusted R^2	0.40	0.23	0.39	0.27	0.28	0.26	0.35
Adjusted R^2 without activity variables	0.28	0.17	0.30	–	0.25	0.13	0.23
Difference in adj R^2	0.12	0.06	0.09	–	0.03	0.13	0.12
"Energy-intensive" activity models							
Adjusted R^2	0.31	0.18	0.41	0.27	0.26	0.28	0.34
Adjusted R^2 without activity variables	0.27	0.15	0.30	–	0.22	0.18	0.23
Difference in adj R^2	0.04	0.03	0.11	–	0.04	0.10	0.11

duration. We expect activities to show increasingly strong associations to more finely resolved electricity readings, and this is something we are exploring in current research.

The second result deserving of further discussion is our finding that the models for the day 1 5–9 p.m. period and the day 2 5–9 p.m. period show considerable differences in the variables selected and in model explanatory power. We note that fewer activities are reported during the day 2 5–9 p.m. period, but given that lasso standardizes predictors prior to selection, this should not influence the algorithm's selection procedure. Furthermore, we note that the 1–5 p.m. model includes several activity variables and a greater explanatory power than the day 2 5–9 p.m. model, even with far fewer activities reported. We take this result to suggest that the relationship between activities and electricity consumption may vary considerably in the same households between days and also that peak hours may be especially variable. This finding, too, warrants further exploration in future research.

Policy implications

These results have implications both for energy demand models and for policy considerations surrounding demand flexibility and DSR. First, we have shown that activity data, whether categorized as "energy-intensive" or not, can improve models of household electricity use. For at least one of our "peak period" models, we find that activity data is especially useful for understanding variations in de-minned electricity use during these hours.

Our novel approach to collect time-use data in parallel with electricity consumption can overcome many of the limitations and challenges previously discussed in the literature on time-use or activity-based models of energy demand (McKenna et al. 2017; Torriti 2014). Specifically, this method enables the collection of more energy-relevant activities, which we have shown to increase the strength and statistical validity of the relationships between household activities and electricity consumption. Given the lack of empirical evidence on these relationships, these results can help improve the specification of activity-appliance signals in models incorporating time-use data. Our approach also facilitates the collection of activity data from multi-occupant households, which have previously been more difficult to account for in time-use models of electricity demand.

Second, while some literature has concluded that it is sufficient to model active occupancy states for the purpose of constructing more accurate energy demand models, we believe that such an approach fails to more directly link the types of activities that are being performed during "active occupancy," which have important consequences and policy implications for delivering more flexible demand.

Our results show that certain types of activities have stronger relationships to electricity consumption at different times of day. Targeting these activities in DSR interventions could yield larger shifts in demand. Furthermore, tailoring DSR interventions based on better evidence about what is actually occurring in households during peak times and about how these activities vary among different segments of the population can improve their effectiveness while also mitigating their impact on vulnerable populations.

In terms of potential for shifting demand alone, these results suggest food preparation and meal times, household chores, and recreational activities should be prioritized for activity-led DSR given their strong associations with electricity use. But this is not the only consideration upon which effective DSR strategies should be based. Other considerations, such as those comprising Torriti et al.'s (2015) "Flexibility Index", are also valuable for determining which activities at which times can be shifted without adverse social impacts.

Study limitations

Some limitations are present in this analysis. The sample is non-representative, and several sample biases are present. The sample size is also small relative to the number of predictors. Standard errors for most variables in the regression models are therefore large in comparison to coefficient size. Models are not constructed for different days of week or seasons in order to preserve the sample size, but these variables are instead included at the model selection stage. Furthermore, self-reported activities can lead to biases and inaccuracies. Measuring activities as simple frequency counts is a coarse way to include these data and may obfuscate more complex relationships between activity type, timing, and electricity consumption. A similar approach as the one taken by Rhodes et al. (2014) and McLoughlin et al. (2015), who cluster load profiles and use household characteristics to build predictive models of membership in distinctive load profiles, may be instructive in this context. Time-use data may be even better suited to this approach than socio-demographic, dwelling, or appliance ownership data, due to its temporal resolution.

Our initial categorization of activity data relies on a taxonomy that was not developed for energy modeling purposes, and although we make an attempt at categorizing activities as "energy-intensive," this is a subjective exercise. We believe much more work could be done to identify more energy-relevant activity categories with useful implications for energy demand modeling and for estimating the potential of demand response (Anderson 2016).

Finally, more data from a more representative sample of households could reduce uncertainty surrounding the strength of relationships between predictive variables and electricity consumption and could make these results more generalizable. The data collection is ongoing, such that more detailed analyses can be conducted in future.

We believe there is potential to scale this research to much larger (> 1000 households), nationally representative samples using app-based activity diaries coupled with smart meter data, and preparations are underway for scaling this study. We also encourage similar studies across varying cultural and societal contexts.

Conclusions

This paper presents analysis of intra-day electricity consumption for a sample of 173 UK households. Electricity provision during peak times is costly and challenging in energy systems with high penetrations of renewable electricity. Understanding the drivers behind electricity consumption during peak times can assist in the design of more effective intervention strategies to shift electricity use patterns in the residential sector.

We present regression models of average full-day and 4-h de-minned electricity consumption using a wide range of predictive factors, including socio-demographic, physical dwelling, appliance ownership, and categorized activity variables. Our analysis tests the ability of these variables to explain consumption at different times of day and the strength of these relationships.

Our results show that adding activity data to regression models can improve their explanatory power of de-minned electricity consumption during most times of day. This result holds regardless of whether only activities that we reasonably assume to be "energy-intensive" are included. Given how challenging it is to achieve greater explanatory power of highly diverse electricity uses in households, a nine-percentage point increase in adjusted R^2 on average is encouraging. The results also show strong associations between activities and electricity consumption at different times of day. The importance of household chores, food, and recreation activities for explaining electricity consumption patterns is clearly demonstrated.

We expect that more nuanced categorizations of activities and investigations of their relationship to electricity consumption at more finely resolved timescales will yield improved insights into the flexibility of demand. This evidence could reduce the cost of integrating renewable energy sources into existing energy systems by providing a better estimate of the potential of demand-side response in the residential sector, thus contributing to accelerating the low-carbon energy transition.

Compliance with ethical standards The University of Oxford's Central University Research Ethics Committee (CUREC) has approved this study.

Conflict of interest The authors declare that they have no conflicts of interest.

Appendix

Table 8 Standardized regression coefficients for Table 5

Predictor	5–9 p.m. (1) β	5–9 p.m. (2) β	9 p.m.–1 a.m. β	1 a.m.–5 a.m. β	5 a.m.–9 a.m. β	9 a.m.–1 p.m. β	1 p.m.–5 p.m. β
Activity categories							
Care for home	0.46***		0.21*		0.22	0.32	0.28
Care for others					0.23	0.49**	
Food	0.16	0.52***	0.25*		0.25		0.31
Other activity							0.37*
Recreation	0.33*		0.25*			0.42*	0.44*
Appliances							
EV				2.23***			
Gas boiler			− 0.37*		− 0.59**		
Heat pump					0.76		
No. of night storage heaters							0.53**
No. of power showers			0.29**			0.35*	
No. of TVs/computers		0.19***				0.46**	0.24
Tumble dryer					0.28		0.37*
Under floor heating			0.11		0.34		
Dwelling							
Detached home	0.62***	0.56**	0.20		0.17		0.66**
No. of rooms in home		0.20**				0.25	0.21
Terraced home	− 0.16						
Socio-demographics							
Bill affordability	− 0.24*						
Estimated monthly electric bill		0.27***	0.18		0.26	0.25	
No. of occupants	0.30*		0.14		0.20		0.29
Owns cat		0.31	0.36**		0.35*		
Participated in summer	− 0.43***						
Participated in spring/autumn			0.34**				
Tariff: Economy 7 or 10				1.73***			
Tariff: Renewable	− 0.42**		− 0.43**				
Tenure (rent vs. own)					− 0.38*		
Constant	6.19	5.66	5.66	4.10	5.28	4.99	4.75
Observations	171	170	172	172	172	170	170
R^2	0.43	0.25	0.43	0.28	0.33	0.29	0.39
Adjusted R^2	0.40	0.23	0.39	0.27	0.28	0.26	0.35

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$

Table 9 Standardized regression coefficients for Table 6

Predictor	5–9 p.m. (1) β	5–9 p.m. (2) β	9 p.m.–1 a.m. β	1 a.m.–5 a.m. β	5 a.m.–9 a.m. β	9 a.m.–1 p.m. β	1 p.m.–5 p.m. β
Activity categories							
Care for home					1.05**	0.44*	0.50**
Care for self					0.21		
Food	0.04	0.43**	0.28**			0.23	0.38*
Recreation	0.40**		0.43***			0.53**	0.44*
Appliances							
EV				2.23***			
Gas boiler			− 0.32*		− 0.65***		
No. of night storage heaters							0.56***
No. of power showers			0.34**			0.29	
No. of TVs/computers						0.38*	0.20
Tumble dryer						0.34	0.38*
Under floor heating			0.12		0.40*	0.38	
Dwelling							
Detached home	0.69***	0.66**	0.20		0.28	0.18	0.76**
No. of rooms in home		0.22				− 0.02	0.19
Terraced home	− 0.15						
Socio-demographics							
Estimated monthly electric bill		0.34*	0.12		0.29	0.13	
No. of occupants	0.39**		0.19		0.35*	0.28	0.35*
Owns cat			0.33**		0.33		
Participated in summer	− 0.43**						
Participated in spring/autumn			0.36**				
Participated on Fri–Sat						0.56*	
Tariff: Economy 7 or 10				1.73***			
Tariff: Renewable	− 0.53**		− 0.40**				
Constant	6.19	5.72	5.62	4.10	5.34	4.69	4.74
Observations	171	170	172	172	172	170	170
R^2	0.43	0.25	0.43	0.28	0.33	0.29	0.39
Adjusted R^2	0.40	0.23	0.39	0.27	0.28	0.26	0.35

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$

References

Anderson, B. (2016). Laundry, energy and time: Insights from 20 years of time-use diary data in the United Kingdom. *Energy Research & Social Science, 22*, 125–136. https://doi.org/10.1016/j.erss.2016.09.004.

Anderson, B., & Torriti, J. (2018). Explaining shifts in UK electricity demand using time use data from 1974 to 2014. *Energy Policy, 123*, 544–557. https://doi.org/10.1016/j.enpol.2018.09.025.

Anderson, B., Lin, S., Newing, A., Bahaj, A., & James, P. (2016). Electricity consumption and household characteristics: implications for census-taking in a smart metered future. *Computers, Environment and Urban Systems, 63*, 58–67. https://doi.org/10.1016/j.compenvurbsys.2016.06.003.

BEIS. (2017). *The clean growth strategy: leading the way to a low carbon future*. Department for Business, Energy & Industrial Strategy.

BEIS. (2018). DUKES chapter 5: statistics on electricity from generation through to sales. In *Digest of UK Energy Statistics (DUKES): electricity*. London: Department for Business,

Energy & Industrial Strategy https://www.gov.uk/government/statistics/electricity-chapter-5-digest-of-united-kingdom-energy-statistics-dukes. Accessed 11 April 2019

CCC. (2016). Annex 2: heat in UK buildings today. In *Next steps for UK heat policy*. London: Committee on Climate Change https://www.theccc.org.uk/wp-content/uploads/2017/01/Annex-2-Heat-in-UK-Buildings-Today-Committee-on-Climate-Change-October-2016.pdf . Accessed 12 February 2019.

De Lauretis, S., Ghersi, F., & Cayla, J.-M. (2017). Energy consumption and activity patterns: an analysis extended to total time and energy use for French households. *Applied Energy, 206*, 634–648. https://doi.org/10.1016/j.apenergy.2017.08.180.

DECC. (2012). *Demand side response in the domestic sector - a literature review of major trials*. London: Department of Energy and Climate Change https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/48552/5756-demand-side-response-in-the-domestic-sector-a-lit.pdf. Accessed 11 April 2019.

DECC. (2013). *Report 9: Domestic appliances, cooking & cooling equipment (No. BRE report number 288143)*. London: Department of Energy and Climate Change https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/274778/9_Domestic_appliances__cooking_and_cooling_equipment.pdf. Accessed 16 March 2018.

DECC. (2014). *Seasonal variations in electricity demand*. London: Department of Energy and Climate Change https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/295225/Seasonal_variations_in_electricity_demand.pdf . Accessed 12 February 2019.

Druckman, A., Buck, I., Hayward, B., & Jackson, T. (2012). Time, gender and carbon: a study of the carbon implications of British adults' use of time. *Ecological Economics, 84*, 153–163. https://doi.org/10.1016/j.ecolecon.2012.09.008.

Elexon. (2013). *Load profiles and their use in electricity settlement* (p. 31). Elexon www.elexon.co.uk/wp-content/uploads/2013/11/load_profiles_v2.0_cgi.pdf.

Ellegård, K., & Palm, J. (2011). Visualizing energy consumption activities as a tool for making everyday life more sustainable. *Applied Energy, 88*(5), 1920–1926. https://doi.org/10.1016/j.apenergy.2010.11.019.

Ellegård, K., Vrotsou, K., & Widén, J. (2010). VISUAL-TimePAcTS/energy use – a software application for visualizing energy use from activities performed. In *Conference proceedings 3rd International Scientific Conference on "Energy Systems with It"* (pp. 1–17). http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.468.5526&rep=rep1&type=pdf.

eurostat. (2009). *Harmonised European time use surveys: 2008 guidelines*. Luxembourg: Office for Official Publications of the European Communities.

Firth, S., Lomas, K., Wright, A., & Wall, R. (2008). Identifying trends in the use of domestic appliances from household electricity consumption measurements. *Energy and Buildings, 40*(5), 926–936. https://doi.org/10.1016/j.enbuild.2007.07.005.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software, 33*(1), 1–22.

Gelman, A. (2008). Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine, 27*(15), 2865–2873. https://doi.org/10.1002/sim.3107.

Gelman, A., Su, Y.-S., Yajima, M., Hill, J., Pittau, M. G., Kerman, J., et al. (2018). *arm: data analysis using regression and multilevel/hierarchical models*. https://CRAN.R-project.org/package=arm.

Gershuny, J. (2008). *Time-use studies: daily life and social change: Full Research Report* (ESRC End of Award Report, RES-000-23-0704-A). Swindon: ESRC. https://s3-eu-west-1.amazonaws.com/esrc-files/outputs/DXDg-pyFREKpOk_6bEon8Q/gWE89bsrZEu6BF9Azmg_Cw.pdf

Gershuny, J., & Sullivan, O. (2017). United Kingdom time use survey, 2014–2015. [data collection]. UK Data Service. SN: 8128. https://doi.org/10.5255/UKDA-SN-8128-1

Grunewald, P., & Diakonova, M. (2018a). Flexibility, dynamism and diversity in energy supply and demand: a critical review. *Energy Research & Social Science, 38*, 58–66. https://doi.org/10.1016/j.erss.2018.01.014.

Grunewald, P., & Diakonova, M. (2018b). The electricity footprint of household activities - implications for demand models. *Energy and Buildings, 174*, 635–641. https://doi.org/10.1016/j.enbuild.2018.06.034.

Grünewald, P., & Diakonova, M. (2019). METER data. Energy-use.org. http://www.energy-use.org/docs/data/. Accessed 11 April 2019.

Grünewald, P., & Layberry, R. (2015). Measuring the relationship between time-use and electricity consumption. In Proceedings of the 2015 ECEEE Summer Study on Energy Efficiency in Buildings (Vol. 9, pp. 2087–96). Presqu'île de Giens, France: ECEEE. www.eceee.org/library/conference_proceedings/eceee_Summer_Studies/2015/9-dynamics-of-consumption/measuring-the-relationship-between-time-use-and-electricity-consumption/.

Grünewald, P., Diakonova, M., Zilli, D., Bernard, J., & Matousek, A. (2017). What we do matters – a time-use app to capture energy relevant activities. In Proceedings of the 2017 ECEEE Summer Study on Energy Efficiency in Buildings (Vol. 9, pp. 2085–93). Presqu'île de Giens, France: ECEEE. www.eceee.org/library/conference_proceedings/eceee_Summer_Studies/2017/9-consumption-and-behaviour/what-we-do-matters-8211-a-time-use-app-to-capture-energy-relevant-activities/.

Harrell, F. (2001). *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis*. New York: Springer-Verlag www.springer.com/gb/book/9781441929181.

Hsu, D. (2015). Identifying key variables and interactions in statistical models of building energy consumption using regularization. *Energy, 83*, 144–155. https://doi.org/10.1016/j.energy.2015.02.008.

Huebner, G., Shipworth, D., Hamilton, I., Chalabi, Z., & Oreszczyn, T. (2016). Understanding electricity consumption: a comparative contribution of building factors, socio-demographics, appliances, behaviours and attitudes. *Applied Energy, 177*, 692–702. https://doi.org/10.1016/j.apenergy.2016.04.075.

Jalas, M., & Juntunen, J. K. (2015). Energy intensive lifestyles: time use, the activity patterns of consumers, and related energy demands in Finland. *Ecological Economics, 113*, 51–59. https://doi.org/10.1016/j.ecolecon.2015.02.016.

Jin, L., Lee, D., Sim, A., Borgeson, S., Wu, K., Spurlock, C. A., & Todd, A. (2017). Comparison of clustering techniques for residential energy behavior using smart meter data (p. 7). Presented at the thirty-first AAAI conference on Artifical intelligence.

Jones, R. V., & Lomas, K. J. (2016). Determinants of high electrical energy demand in UK homes: appliance ownership and use. *Energy and Buildings, 117*, 71–82. https://doi.org/10.1016/j.enbuild.2016.02.020.

Jones, R. V., Fuertes, A., & Lomas, K. J. (2015). The socioeconomic, dwelling and appliance related factors affecting electricity consumption in domestic buildings. *Renewable and Sustainable Energy Reviews, 43*, 901–917. https://doi.org/10.1016/j.rser.2014.11.084.

Kavousian, A., Rajagopal, R., & Fischer, M. (2013). Determinants of residential electricity consumption: using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behavior. *Energy, 55*, 184–194. https://doi.org/10.1016/j.energy.2013.03.086.

Lader, D., Short, S., & Gershuny, J. (2006). *The time use survey, 2005* (p. 94). London: Office for National Statistics www.timeuse.org/sites/ctur/files/public/ctur_report/1905/lader_short_and_gershuny_2005_kight_diary.pdf.

McKenna, E., & Thomson, M. (2016). High-resolution stochastic integrated thermal–electrical domestic demand model. *Applied Energy, 165*, 445–461. https://doi.org/10.1016/j.apenergy.2015.12.089.

McKenna, E., Higginson, S., Grunewald, P., & Darby, S. J. (2017). Simulating residential demand response: improving socio-technical assumptions in activity-based models of energy demand. *Energy Efficiency, 11*, 1583–1597. https://doi.org/10.1007/s12053-017-9525-4.

McLoughlin, F., Duffy, A., & Conlon, M. (2012). Characterising domestic electricity consumption patterns by dwelling and occupant socio-economic variables: an Irish case study. *Energy and Buildings, 48*, 240–248. https://doi.org/10.1016/j.enbuild.2012.01.037.

McLoughlin, F., Duffy, A., & Conlon, M. (2015). A clustering approach to domestic electricity load profile characterisation using smart metering data. *Applied Energy, 141*, 190–199. https://doi.org/10.1016/j.apenergy.2014.12.039.

National Infrastructure Commission. (2016). *Smart Power* (National Infrastructure Commission Report). https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/505218/IC_Energy_Report_web.pdf. Accessed 18 March 2018.

Ofgem. (2010). *Demand side response. A discussion paper* (p. 87). London: Office of Gas and Electricity Markets https://www.ofgem.gov.uk/ofgem-publications/57026/dsr-150710.pdf. Accessed 11 April 2019.

Ofgem. (2017). *Typical domestic consumption values for gas and electricity.* London: Office of Gas and Electricity Markets https://www.ofgem.gov.uk/system/files/docs/2017/08/tdcvs_2017_open_letter.pdf . Accessed 20 February 2019.

ONS. (2016). Population estimates for UK, England and Wales, Scotland and Northern Ireland: mid-2015. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2015. Accessed 29 February 2019.

ONS. (2017). Families and households: 2017. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsand?marriages/families/bulletins/familiesandhouseholds/2017. Accessed 26 February 2019.

ONS. (2018a). English housing survey: headline report, 2016–17. Ministry of Housing, Communities & Local Government. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/675942/2016-17_EHS_Headline_Report.pdf. Accessed 26 February 2019.

ONS. (2018b). Percentage of households with durable goods, UK: Table A45. https://www.ons.gov.uk/peoplepopulation?andcommunity/personalandhouseholdfinances?/expenditure/datasets/percentageofhouseholds?withdurablegoodsuktablea45. Accessed 26 February 2019.

Palmer, J., Terry, N., Kane, T., Firth, S., Hughes, M., Pope, P., et al. (2013). Further analysis of the household electricity use survey - electrical appliances at home: tuning in to energy saving (No. 475/09/2012). Cambridge Architectural Research Limited, element energy, and Loughborough University. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/275484/electricity_survey_2_tuning_in_to_energy_saving.pdf.

Powells, G., Bulkeley, H., Bell, S., & Judson, E. (2014). Peak electricity demand and the flexibility of everyday life. *Geoforum, 55*, 43–52. https://doi.org/10.1016/j.geoforum.2014.04.014.

R Development Core Team. (2008). *R: a language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing www.R-project.org.

Rhodes, J. D., Cole, W. J., Upshaw, C. R., Edgar, T. F., & Webber, M. E. (2014). Clustering analysis of residential electricity demand profiles. *Applied Energy, 135*, 461–471. https://doi.org/10.1016/j.apenergy.2014.08.111.

Richardson, I., Thomson, M., Infield, D., & Clifford, C. (2010). Domestic electricity use: a high-resolution energy demand model. *Energy and Buildings, 42*(10), 1878–1887. https://doi.org/10.1016/j.enbuild.2010.05.023.

Roberts, N., & Thatcher, J. (2009). Conceptualizing and testing formative constructs: tutorial and annotated example. *ACM SIGMIS Database, 40*(3), 9–39.

Satre-Meloy, A. (2019). Investigating structural and occupant drivers of annual residential electricity consumption using regularization in regression models. *Energy, 174*, 148–168. https://doi.org/10.1016/j.energy.2019.01.157.

Sekar, A., Williams, E., & Chen, R. (2018). Changes in time use and their effect on energy consumption in the United States. *Joule., 2*, 521–536. https://doi.org/10.1016/j.joule.2018.01.003.

Stankovic, L., Stankovic, V., Liao, J., & Wilson, C. (2016). Measuring the energy intensity of domestic activities from smart meter data. *Applied Energy, 183*, 1565–1580. https://doi.org/10.1016/j.apenergy.2016.09.087.

Strbac, G., Aunedi, M., Pudjianto, D., Djapic, P., Teng, F., Sturt, A., et al. (2012). *Strategic assessment of the role and value of energy storage systems in the UK low carbon energy future.* London: Imperial College http://www.efn-uk.org/l-street/politics-lib/gov-agents/index_files/ICLenergy-storage2012.pdf . Accessed 24 February 2019.

Swan, L. G., & Ugursal, V. I. (2009). Modeling of end-use energy consumption in the residential sector: a review of modeling

techniques. *Renewable and Sustainable Energy Reviews, 13*(8), 1819–1835. https://doi.org/10.1016/j.rser.2008.09.033.

Tabbone, L., Ravalet, E., Durand-Daubin, M., Kaufmann, V., & Cayla, J.-M. (2016). Spatial location of activities and energy consumption of households in France. Presented at the DEMAND Centre Conference, Lancaster. http://www.demand.ac.uk/wp-content/uploads/2016/05/DEMAND2016_Full_paper_131-Tabbone-v2.pdf. Accessed 11 April 2019.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B: Methodological, 58*, 267–288.

Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society, Series B: Statistical Methodology, 73*(3), 273–282.

Torriti, J. (2014). A review of time use models of residential electricity demand. *Renewable and Sustainable Energy Reviews, 37*, 265–272. https://doi.org/10.1016/j.rser.2014.05.034.

Torriti, J., Hanna, R., Anderson, B., Yeboah, G., & Druckman, A. (2015). Peak residential electricity demand and social practices: deriving flexibility and greenhouse gas intensities from time use and locational data. *Indoor and Built Environment, 24*(7), 891–912. https://doi.org/10.1177/1420326X15600776.

U.S. Bureau of Labor Statistics. (2018). *American time use survey: handbook of methods*. Washington, D.C.: U.S. Department of Labor https://www.bls.gov/opub/hom/atus/pdf/atus.pdf . Accessed 26 February 2019.

Viegas, J. L., Vieira, S. M., Melício, R., Mendes, V. M. F., & Sousa, J. M. C. (2016). Classification of new electricity customers based on surveys and smart metering data. *Energy, 107*, 804–817. https://doi.org/10.1016/j.energy.2016.04.065.

Wickham, H. (2016). *ggplot2: elegant graphics for data analysis*. New York: Springer-Verlag http://ggplot2.org.

Wickham, H., François, R., Henry, L., & Müller, K. (2019). *dplyr: a grammar of data manipulation*. https://CRAN.R-project.org/package=dplyr.

Widén, J., & Wäckelgård, E. (2010). A high-resolution stochastic model of domestic activity patterns and electricity demand. *Applied Energy, 87*(6), 1880–1892. https://doi.org/10.1016/j.apenergy.2009.11.006.

Widén, J., Lundh, M., Vassileva, I., Dahlquist, E., Ellegård, K., & Wäckelgård, E. (2009). Constructing load profiles for household electricity and hot water from time-use data—modelling approach and validation. *Energy and Buildings, 41*(7), 753–768. https://doi.org/10.1016/j.enbuild.2009.02.013.

Wyatt, P. (2013). A dwelling-level investigation into the physical and socio-economic drivers of domestic energy consumption in England. *Energy Policy, 60*, 540–549. https://doi.org/10.1016/j.enpol.2013.05.037.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B: Statistical Methodology, 67*(2), 301–320.