WWW.FRONTIER-ECONOMICS.COM

PROJECT VENICE: THE IMPACT OF THE PANDEMIC ON ELECTRICITY CONSUMPTION

20 MARCH 2023

CONTENTS

Note: Neither the European Commission nor the European Centre for Medium-Range Weather Forecasts is responsible for any use that may be made of the Copernicus information *or data it contains.*

This report was co-produced by Frontier Economics and University College London (Dr Ellen Zapata-Webborn and Dr Eoghan McKenna of UCL). Please cite as:

Frontier Economics, Zapata-Webborn, E., McKenna, E. (2023) Project VENICE: The impact of the pandemic on electricity consumption.

Executive summary

Frontier Economics was commissioned by National Grid Electricity Distribution (NGED) as part of Project VENICE (Vulnerability and Energy Networks, Identification and Consumption Evaluation) to analyse the potential impact that the COVID-19 pandemic may have had on electricity networks, and how the impact might persist over the ED2 period.

We considered the following questions in this analysis.

- How did households respond to the pandemic in terms of their electricity consumption?
- How might these changes in electricity consumption persist?
- How did these changes, and their persistence, vary by customer attribute (age, employment status, etc.)?
- What are the implications for NGED's network planning?
- What are the implications for vulnerable customers connected to NGED's networks?

To answer these questions, we undertook the following analysis:

- We first developed hypotheses on changes in electricity consumption in response to the pandemic and the likely persistence of these changes.
- We then tested the hypotheses through analysis of smart meter data.

Hypotheses

We generated hypotheses, based on a review of the literature and emerging published data, and the application of a behavioural framework around habit formation. This review found that:

- During the pandemic, specific customer groups stayed at home more and therefore increased electricity consumption (office workers working from home, households with school aged children, furloughed workers, those who became unemployed during the pandemic, and elderly and clinically vulnerable people). Each of these groups tends to have different consumption patterns, and therefore the impact of their change in behaviour varies. Overall, the review found the following hypotheses:
    - the pandemic was likely to have increased domestic electricity consumption during the day; and
    - the pandemic had a smaller impact in the evenings.
- Applying our behavioural framework suggested a difference in how these changes may have persisted across groups. Our assessment suggested that behaviours with high net rewards (such as working from home) were likely to persist. Other behaviours, such as those driven by being furloughed or unemployed, were less likely to persist.

Data analysis

We then tested these hypotheses through data analysis.

Individual households

We first analysed the impact of the pandemic on individual household electricity consumption. We used smart meter data from the Smart Energy Research Lab, which contains half-hourly data consumption over the pandemic period for just under 600 households. This included survey data that we could use to identify household characteristics. We then used pre-pandemic data to estimate counterfactual electricity consumption patterns, which represented estimated consumption had the pandemic not occurred. We used two different predictive modelling approaches (a linear regression and machine learning) which modelled each household separately. Overall, we found that actual electricity consumption increased on average versus predicted consumption; which can be attributed to the impact of the pandemic. More specifically, the greatest increases took place during working hours, and the greatest changes in consumption occurred during lockdown periods. And whilst on average the effect of the pandemic tended to diminish over the course of the pandemic period, it did not fully diminish for all households.

We investigated the changes in consumption for different sub-groups of customers in two ways.

- First we picked characteristics of customers (based on responses to survey data) and analysed changes in consumption for those customers.
- Then we used a clustering algorithm to group customers that showed similar behaviours in electricity consumption.

We found that there are actually many different responses to the pandemic within a typical day and across the entire period for different types of customers. Analysing customers based on their characteristics meant it was possible to observe differences in response to the pandemic, and we were able to test some of our more specific hypotheses (subject to data limitations). However, our clustering analysis allowed us to identify specific behavioural groups independent of customer characteristics; illustrating the different types of responses.

Finally, we looked at the impact on NGED, considering the impact on vulnerable customers and the network:

- Vulnerable customers tended to have lower changes in consumption than other types of customers, and hence there was a small bill impact.
- Only a small proportion of NGED’s network appears to be at risk of constraints if an increase in peak demand seen during the pandemic continues into the future.

Persistence

The second part of our analysis considered how behaviours might persist. In terms of future electricity consumption, we found that there are only a small subset of customers that are likely to have a sustained increase to their electricity consumption. This was based on our behavioural framework which analysed different groups of households and attempted to understand which are most likely to receive net rewards as a result of their changing behaviour. We triangulated this with evidence from our modelling to narrow down households that will remain persistent in their behaviour.

The most prominent such group are households that continue working from home. Within this group there are those who are likely to receive high net rewards from working from home, including households with children. We found that these customers had higher electricity consumption during the second half of 2021 compared to other households. However, that increased consumption is likely to be small and isolated to daytime periods.

There is the potential for this to impact a geographically small part of NGED’s network that are close to facing capacity constraints. While these trends are small, future electrification of households (through increased uptake of heat pumps and electric vehicles) could increase consumption further for these customers if their patterns of home working persist.

We note that the “cost of living crisis”, associated with high energy costs and economic slowdown in 2022, may have led to customers changing behaviour again. In particular, it might have led to customers decreasing their consumption.

1 Context, introduction and aims

Frontier Economics was commissioned by National Grid Electricity Distribution (NGED) as part of Project VENICE (Vulnerability and Energy Networks, Identification and Consumption Evaluation) to analyse the potential that the COVID-19 pandemic may have on electricity networks in the short to medium term.

On 23 March 2020, the British Government issued a stay-at-home order in response to the COVID-19 pandemic, otherwise known as "lockdown". As a result, all non-essential shops and services were closed and people were instructed to stay home, other than for essential trips. For the next two years, there were some form of policy measures or guidance relating to the pandemic. A summary of the key measures, as based on Institute for Government analysis[1], are:

- In March 2020, the first national lockdown was introduced, which lasted until late June 2020.
- Throughout the summer and autumn of 2020 a variety of measures were introduced until, in November 2020, a second national lockdown was introduced. This lasted until early December, where other regional measures and restrictions were introduced.
- In early January 2021, a third national lockdown was introduced. This remained in force for two months, after which point the restrictions were reduced through to the summer.
- Legal restrictions were mostly lifted by July 2021. The spread of the Omicron variant in winter 2021 did not lead to significant national restrictions. At the time of writing (December 2022), there had been no further national lockdowns in the UK.

As a result of the pandemic and the associated restrictions, people tended to stay at home more over the period from March 2020 to December 2021.

NGED are interested in the behaviours that changed during the pandemic, how these affected electricity consumption and whether they are likely to continue to affect electricity consumption in the future. To address these issues, we have considered three questions:

- What happened to domestic electricity consumption during the pandemic, and how did it differ across customers?
- To what extent might observed changes in consumption during the pandemic persist into ED2, and how does this vary by customer type?
- How will these changes impact NGED's network and vulnerable customers?

To answer these questions, we have structured the report in three sections:

- **Section 2 - Hypothesis generation**. We first generate hypotheses on behaviour change during the pandemic, based on a review of the literature and available data. Within this section we have set out

[1] Institute for Government analysis, https://www.instituteforgovernment.org.uk/charts/uk-government-coronavirus-lockdowns

what customer attributes might affect electricity consumption, how behaviours for those with specific customer attributes changed during the pandemic, and what that means for electricity consumption. We then apply a framework on habit formation to assess how likely these behaviours may be to persist. The outcome of this analysis is a set of hypotheses for testing.

- **Section 3 - Data analysis to test hypotheses.** We then test our hypotheses, through analysis of smart meter data. Within this section we consider the overall impact of the pandemic on household consumption, which includes the aggregate impact across the whole period, as well as considering what happened in specific sub-daily periods. We then investigate the impact on specific customer types in more detail, which includes looking at sub-groups of customers and clustering customers into groups based on their consumption behaviour.

- **Section 4 - Conclusions.** Finally, we have a section that summarises how we expect the effect of the pandemic to persist, and what behaviours developed during the pandemic will continue. We also consider what it means in the future for NGED in terms of its network and for vulnerable customers.

Alongside these sections, we have also included a set of detailed Annexes.

2 Hypotheses for consumption changes during the pandemic

2.1 Introduction

This section sets out our process for generating the hypotheses which we then test in the data analysis. There are three parts to our hypothesis generation.

- We first review the literature and available data from the pandemic, to assess the key customer attributes that may have driven changes in domestic electricity consumption during the pandemic.
- We then assess the likelihood of this behaviour persisting, using a behavioural framework for habit formation. We present an example of our application of this framework (on working from home), with the full set of analysis included in Annex 1 – Customer behaviour framework for habit formation.
- Finally, we summarise hypotheses on changes to behaviour by customer group, and their likelihood of persisting.

2.2 Identifying key customer attributes

To consider how consumption may vary for specific sub-groups of customers, we draw on the literature and available data to investigate:

- Which customer attributes may affect electricity consumption?
- How have behaviours changed during the pandemic and what does this mean for electricity consumption?

2.2.1 Customer attributes and electricity consumption

Table 1 summarises a review of the literature on how electricity use varies by customer attribute. This review was carried out in 2021, and covers the UK (Northern Ireland), as well as the USA and European countries such as Ireland, Denmark and Portugal.[2] [3]

The literature in this area tends to feature relatively basic analysis, particularly when considering the consumption in the counterfactual. When comparing customer consumption profiles, most studies do a simple comparison to the same period in the previous year. There is significant scope to do more sophisticated analysis to identify the impact of changes in customer the analysis in this paper is much more sophisticated, and more likely to isolate changes in customer behaviour.

---

[2] We used the following search terms in Google Scholar: “Energy behaviours”, “Energy behaviours Covid 19”, “Energy behaviour vulnerable households”. We looked at the top 100 results, ruling out studies that were from countries with very different economies and climates, and more than 15 years old.

[3] Much of the remaining literature has studied electricity consumption at a more aggregated level, specifically at distribution or transmission network level. At this level, many of the studies found an overall reduction in consumption, primarily driven by a decline in commercial and industrial consumption rather than domestic consumption. This decline has been cited as between 5% and 10% on a global basis (International Energy Agency, https://www.iea.org/reports/covid-19-impact-on-electricity) to as much as 30% on a domestic basis (ElectraLink, https://www.electralink.co.uk/2022/01/covid-change-elec-consumption-productivity/)

Table 1 Summary of literature review

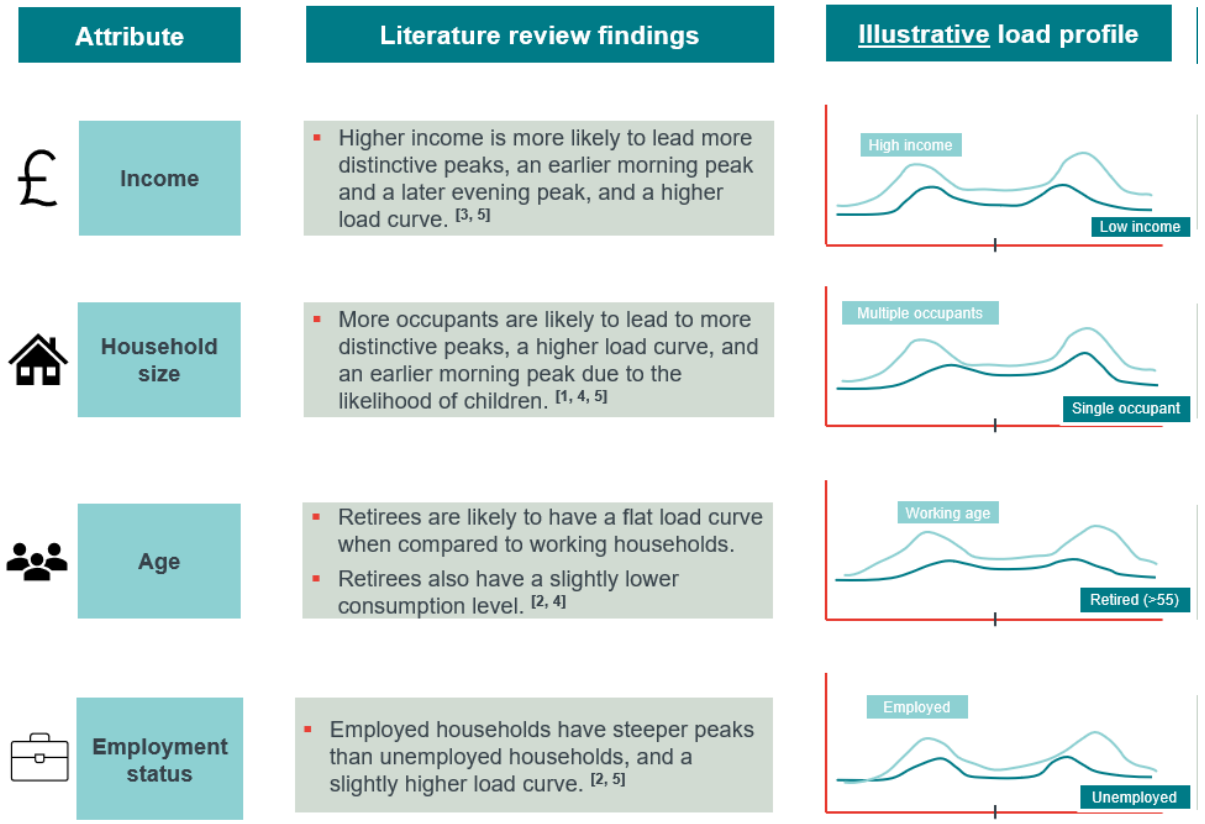
Title	Authors	Geography	Findings
Electricity consumption and household characteristics - Implications for census-taking in a smart metered future	Anderson, B., et al. (2016),	Ireland	Morning peak: Earliest peak is for employed, while largest peak is for carers. Retired have the lowest peak, while unemployed also have a relatively small peak. Evening peak: Employed have the biggest peak, while retired and unemployed have smaller peaks. Carers have the latest peak.
Real life energy use in the UK: How occupancy and dwelling characteristics affect domestic electricity use	Yohanis, Y.G., et al (2008)	Northern Ireland	Income: Upper tercile has earliest and largest morning peak, as well as latest and largest evening peak. Middle tercile morning peak is almost as large, but occurs several hours later, while 'morning' peak for lowest tercile is small and occurs slightly later. Middle tercile evening peak occurs, same time as lowest tercile peak, but lowest tercile peak is notably smaller.
Association Rule Mining Based Quantitative Analysis Approach of Household Characteristics Impacts on Residential Electricity Consumption Patterns	Wang, F., et al (2018)	Ireland	Older HHs, retired people and single occupants are more likely to have flat usage with no distinct peak across the 24hr period. HHs with several people, i.e. working parents with children and/or 3-4 bedroom HHs, have more distinct morning and evening peaks. Presence of children may lead to earlier morning peaks.
An empirical analysis of domestic electricity load profiles: Who consumes how much and when	Trotta, G. (2020)	Denmark	Several occupants (>= 4 people) is associated with higher level of consumption across 24 hr period, as well as more distinctive peaks. These HHs are likely to have >= 1 child, and so morning peaks are likely to be relatively earlier.

Title	Authors	Geography	Findings
			Higher household income is associated with a clear morning peak and a larger distinctive evening peak, whereas lower income households have smaller peaks and a load pattern at a lower level of consumption.
A Clustering Approach to Domestic Electricity Load Profile Characterisation Using Smart Metering Data	McLoughlin F., et al. (2014)	Ireland	More occupants are likely to lead to more distinctive peaks, a higher load curve, and an earlier morning peak due to the likelihood of children. Single occupants are more likely to have flat usage with no distinct peak across the 24hr period.
Daily electricity consumption profiles from smart meters: Proxies of behaviour for space heating and cooling	Gouveia, J. et al. (2014)	Portugal	When comparing the load profiles for active and non-active clusters, the active cluster (higher load profile and more distinctive peaks) had a higher proportion of high earners when compared to the non-active cluster.

Source: Frontier Economics

This evidence suggests four observable customer attributes tend to impact the level and shape of electricity consumption: income, household size, age and employment status. The findings are summarised in Figure 1.

Figure 1 Electricity consumption and key customer attributes



Source: Frontier Economics

2.2.2 The impact of COVID-19

We now consider how the impacts of COVID-19 may have varied for customers with different attributes in these categories.

Our evidence review was carried out in summer 2021 when COVID-19 lockdowns in some cases were still ongoing (see Table 2 for a summary). Based on this, we found that COVID-19 led to two categories of changes in customers' behaviour that could impact on electricity demand.

- As would be expected, there is clear evidence that customers stayed at home more during the pandemic.
- There is also evidence that household size changed during the pandemic, as people moved their primary residences.

We provide more detail on both of these below.

Table 2 Summary of evidence on the impact of COVID-19 on customer behaviour

Title	Authors	Geography	Findings
Why working from home will stick	Barrero,et al (2020)	USA	20% of full workdays will be supplied from home after the pandemic ends, compared with just 5% before.
Changes in Electricity Load Profiles Under COVID-19: Implications of "The New Normal" for Electricity Demand	Burleyson, C., et al. (2020)	USA	COVID-19 shifted weekday load profiles to closely resemble weekend profiles from previous years. They find that long-term structural changes to the workplace like widespread teleworking could lead to 5-7% higher spring and summertime peak hourly loads occurring up to 2.5 hours earlier. Residential weekday load profile: dramatic change in shape, with a more gradual morning ramp, higher midday loads, and a smaller and less steep ramp to the evening peak.
Household composition, couples' relationship quality, and social support during lockdown	Zilanawala, A., et al (2020)	UK	About two-fifths of 19-year-olds, who were living independently of their parents prior to the COVID-19 outbreak, moved in with their parents (or parents-in-law).
Parenting in lockdown: Coronavirus and the effects on work-life balance	ONS (2020)	UK	Parents changed their routines to accommodate their new childcare commitments. Those parents who worked outside of the home contributed childcare outside the usual 'nine to five' hours, which suggests that they were more likely to be working unsociable hours. Unsociable hours for those working away from home would see their electricity loads

Title	Authors	Geography	Findings
			spread to earlier and later in the day, with a more steady increase and fall at the beginning and end of the day.
Living longer: older workers during the coronavirus (COVID-19) pandemic	ONS (2021)	UK	Those employees aged 50 years and over were more likely to report working fewer hours than usual than those under 50. Those aged 65 years and over the most likely to say they had worked reduced hours, they're also the most likely to receive no pay and the least likely to receive full pay. Older people who become unemployed are more likely to be at risk of long-term unemployment than younger people.
Personal and economic well-being in Great Britain: May 2021	ONS (2021)	UK	Groups of people financially impacted at start of pandemic were still worse of as of mid-April 2021, such as self-employed. Lowest income bracket continued to be more likely to report negative impacts to personal well-being in comparison with higher brackets. Those aged under 30 years were consistently more likely to report that their income had been reduced than those over 60 years.
Coronavirus comes home? Energy use, home energy management, and the social-psychological factors of COVID-19	Chen, C. et al (2020)	USA	COVID-19 has led to morning and evening peaks disappearing. Half of respondents reported a higher volume of electricity than before, while only a few reported lower usages.

Title	Authors	Geography	Findings
Electricity demand during pandemic times: The case of the COVID-19 in Spain	Santiago, I. et al (2020)	Spain	COVID-19 has led to lower electricity consumption, resulting in a lower level of the load profile. The morning peak now occurs later, while the evening peak is later and smaller.
Changes in Domestic Energy and Water Usage during the UK COVID-19 Lockdown Using High-Resolution Temporal Data	Meneer, D. et al (2021)	UK	Morning electricity usage occurred later in the day, with usage decreasing from 6am and increasing from midday.
Review analysis of COVID-19 impact on electricity demand for residential buildings.	Krarti, M. and Aldubyan, M. (2021)	Multiple countries	Residential electricity consumption increased by 11% to 32% across several countries during the lockdown of spring 2020.
The impact of the COVID-19 on households' hourly electricity consumption in Canada	Abdeen, A. et al (2021)	Canada	Analysis of the change in electricity usage consumption for 500 homes in Ottawa, Canada, found that average household daily electricity consumption increased by about 12% in 2020 relative to 2019. In addition, the five largest post-COVID-19 (July 2020) peak loads are around 15% to 20% higher than the five largest peak pre-COVID-19 (July 2019) peak loads.

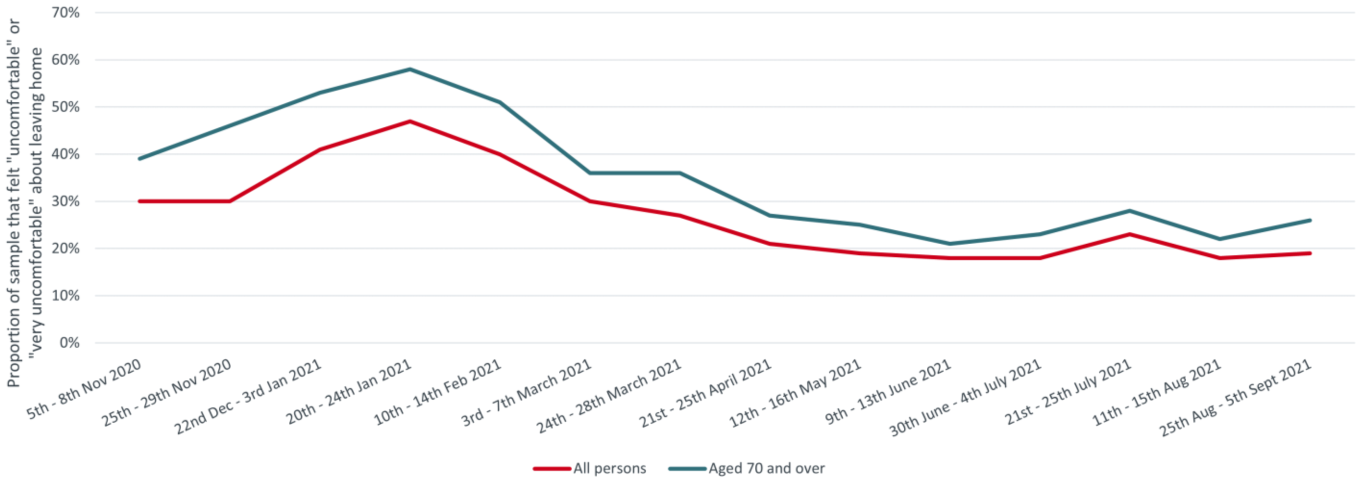
Source: Frontier Economics

Evidence that customers stayed at home

The data suggests that customers stayed at home, for the following reasons.

- **Working from home increased.** 12% of respondents to the Annual Population Survey said they worked from home during 2019, however the proportion of people working from home according to the ONS homeworking survey in April 2020 was 45%[4].
- **Shielding from risk is likely to have led some populations to stay at home more.** Elderly people may have stayed at home more than the general population, due to the increased health risks for the elderly associated with COVID-19. ONS data suggests that those aged over 70 were more likely to feel uncomfortable about leaving the house during the pandemic, but this effect tended to subside over time (Figure 2).

Figure 2 The elderly tended to feel more uncomfortable about leaving home than others during the pandemic



Source: ONS, Coronavirus and the social impacts on Great Britain, Table 8 https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandwellbeing/datasets/coronavirusandthesocialimpactsongreatbritaindata/current

Note: During the lockdown periods, elderly people were more likely to feel uncomfortable about leaving their home

- **School-age children spent more time at home.** The proportion of parents who said a child in their household had been home-schooled between 7th May and 7th June 2020 because of COVID-19 was 87%[5]. The proportion of children home-schooled prior to COVID-19 was around 0.6%[6].
- **The unemployment rate rose**. An increase in the unemployment rate is likely to have resulted in more people spending time at home (Figure 3). The unemployment rate was around 4% across 2019, whilst we can see that it increases during the second half of 2020, with a rate of over 5% in November and December. The rate then gradually falls across 2021, with a rate of 4% in December 2021.

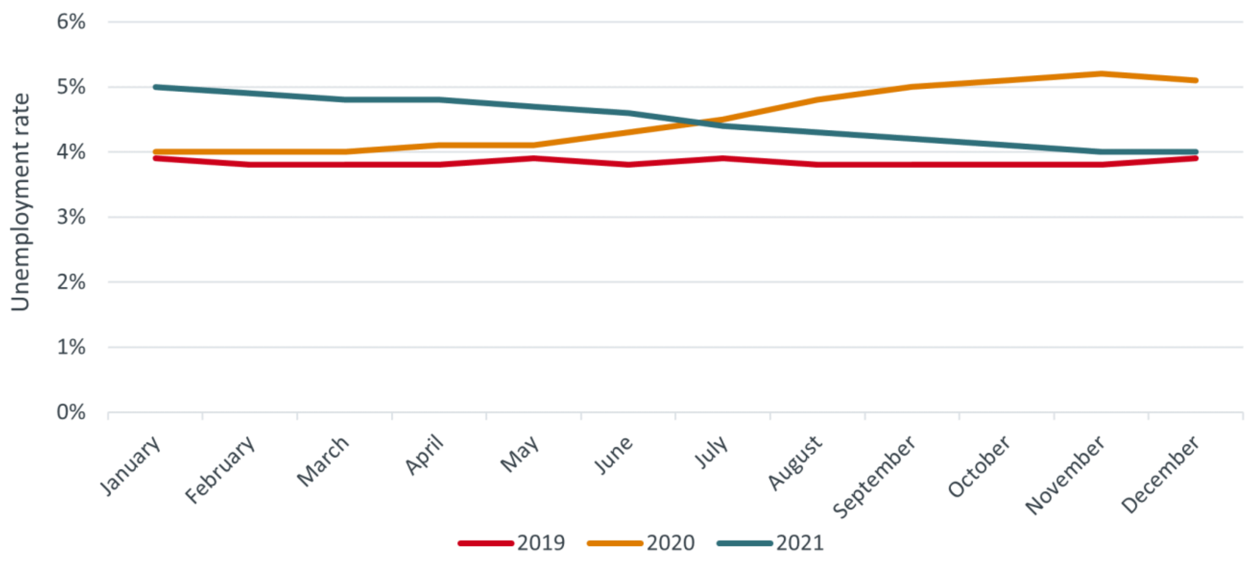
---

[4] ONS, *Annual Population Survey, 2020*, *https://www.nomisweb.co.uk/datasets/apsnew; ONS, Coronavirus and homeworking in the UK: April 2020, https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/bulletins/coronavirusandhomeworkingintheuk/april2020*

[5] ONS, Coronavirus and homeschooling in Great Britain: April to June 2020, https://www.ons.gov.uk/peoplepopulationandcommunity/educationandchildcare/articles/coronavirusandhomeschoolingingreatbritain/apriltojune2020

[6] Number for England, based on: Department for Education, Schools, pupils and their characteristics: January 2019; https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/812539/Schools_Pupils_and_their_Characteristics_2019_Main_Text.pdf; House of Commons Library, Home Education in England, https://commonslibrary.parliament.uk/research-briefings/sn05108/#:~:text=In%202019%2C%20all%20local%20authorities,England%20as%20of%20March%202019.

Figure 3 The unemployment rate increased during 2020



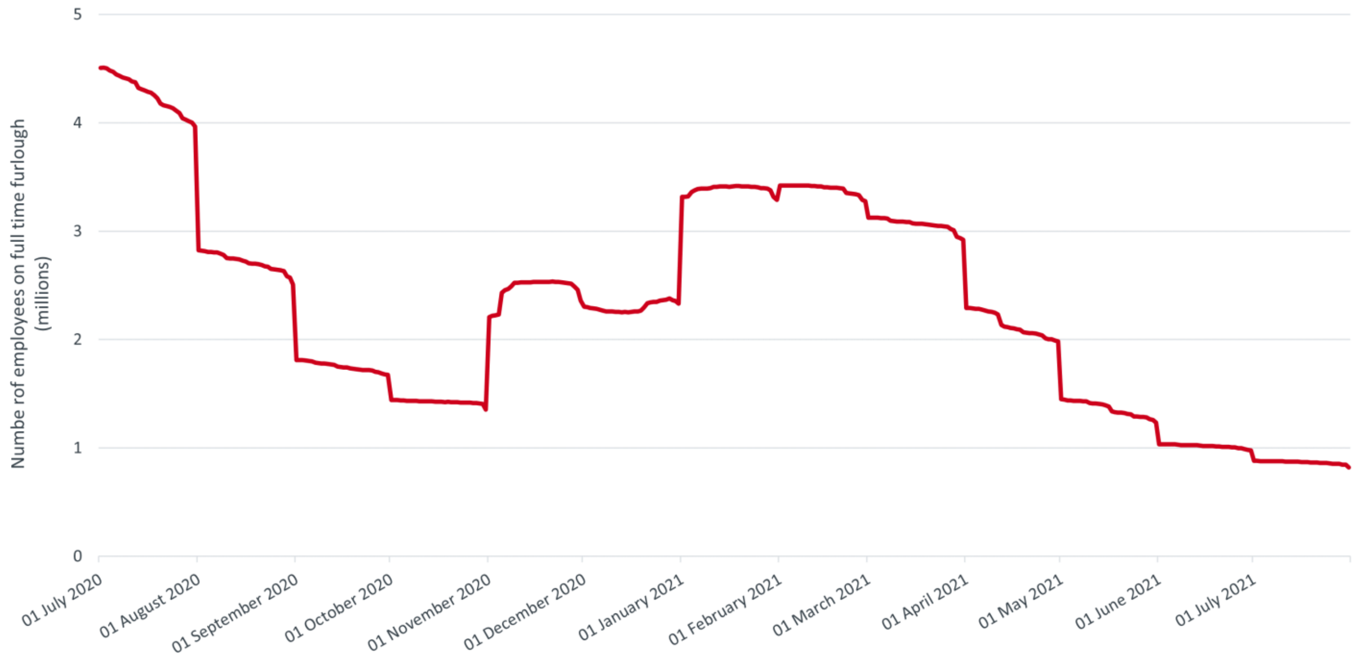
Source: ONS Labour Market Survey, 2022, [*https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/datasets/labourmarketstatistics*](https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/datasets/labourmarketstatistics)

- **A significant number of people, previously in work, were furloughed.** The Government's Job Retention Scheme allowed employers to claim funding for furloughed employees[7]. Furloughed employees were not permitted to work during their period of furlough[8]. This led to more people staying at home.

[7] https://www.gov.uk/guidance/claim-for-wage-costs-through-the-coronavirus-job-retention-scheme

[8] It was possible to flexibly furlough employees. Flexibly furloughed employees could work for any amount of time, and any work pattern but they could not work during hours that you record them as being on furlough. https://www.gov.uk/guidance/claim-for-wage-costs-through-the-coronavirus-job-retention-scheme

Figure 4 Furloughed workers may also have spent more time at home



Source: HMRC, 2021

Changes to the size of households

We also found evidence that household size changed.

- **More young people moved home**. For example, 42% of 19 year olds moved in with their parents (or parents-in-law), having lived independently of their parents prior to the COVID-19 outbreak[9].
- **A significant number of non-UK born people returned to their home countries.** In addition, there was an increase in the number of non-UK nationals leaving the UK, with 0.9m - 1.3m leaving between Q3 2019 and Q3 2020[10].

2.2.3 Implications

Figure 5 draws together the findings of our review of the customer attributes that drive electricity consumption, and the changes that were experienced during COVID-19.

[9] *Zilanawala et al., 2020*. Household composition, couples' relationship quality, and social support during lockdown – Initial findings from the COVID-19 Survey in Five National Longitudinal Studies. London: UCL Centre for Longitudinal Studies.

[10] *Economic Statistics Centre of Excellence, 2021.* 'Estimating the UK population during the pandemic'. Available at: https://www.escoe.ac.uk/estimating-the-uk-population-during-the-pandemic/

Figure 5 Behaviour change and the pandemic: summary of findings from the literature review

There are some behavioural changes that have occurred during the pandemic…	…which have an impact on some key customer attributes
Increased working from home	Employed workers may have increased electricity consumption and higher peaks
Elderly and clinically vulnerable shielding	These customers have flatter daily demand patterns compared to the average. An increase across the day may be observed
School-age children spending more time at home	This may impact on electricity consumption at home during school hours.
Larger number of unemployed people spending time at home	Unemployed households tend to have lower consumption and less steep peaks, so this may effect both total consumption and shape.
More people staying at home due to furlough scheme	Flatter consumption may be expected, when households spend more time at home (but are not necessarily experiencing major falls in income.
Increase in the number of young people moving back home	Families may have marginally increased consumption
Increase in the number of non-UK nationals leaving the UK	Reduced population size may lead to overall reduced electricity consumption

Source: Frontier Economics

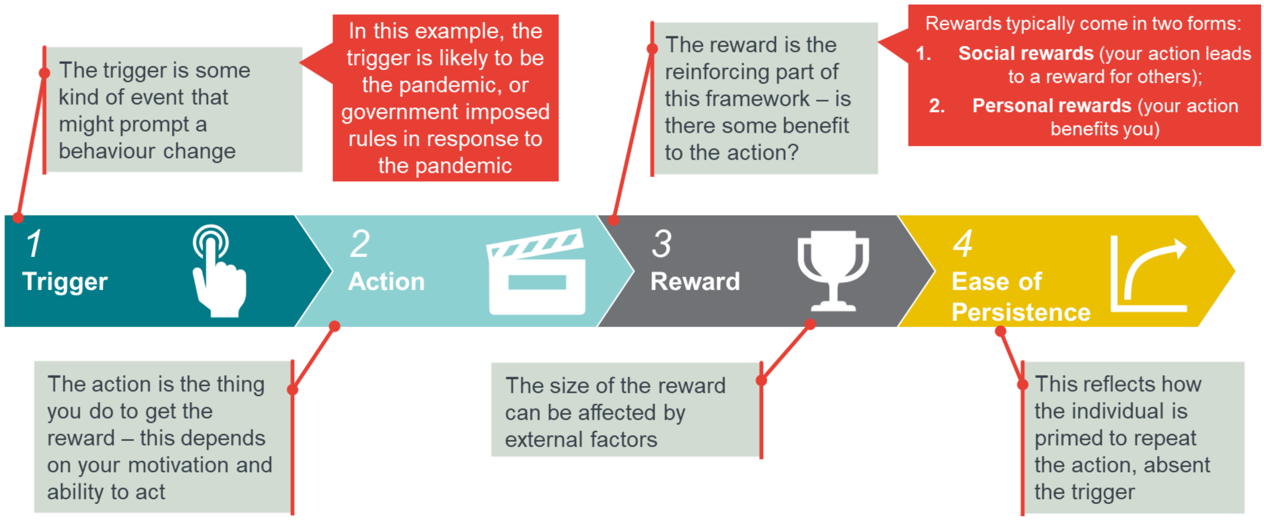
In the next section, we develop hypotheses on whether these change may persist.

2.3 Persistence assessment

To assess persistence, we use a framework around habit formation[11] (Figure 6). In this framework, a trigger can cause an action. If this action has some kind of reward, that will mean individuals are more likely to repeat the action, particularly if some "efforts" to perform the action first time round don't need to be repeated in order to continue the behaviour. We apply this framework to the pandemic, thinking about how the pandemic acted as a trigger for new behaviour and how repeated actions during the pandemic might lead to new behaviours becoming ingrained habits over the course of ED2.

[11] This framework builds on the literature around habit formation, including the Hook Model of Behavioural Design, described for example in this study: Filippou, J. et al (2015), Combining The Fogg Behavioural Model And Hook Model To Design Features In A Persuasive App To Improve Study Habits, https://www.researchgate.net/publication/303223616_Combining_The_Fogg_Behavioural_Model_And_Hook_Model_To_Design_Features_In_A_Persuasive_App_To_Improve_Study_Habits/citation/download

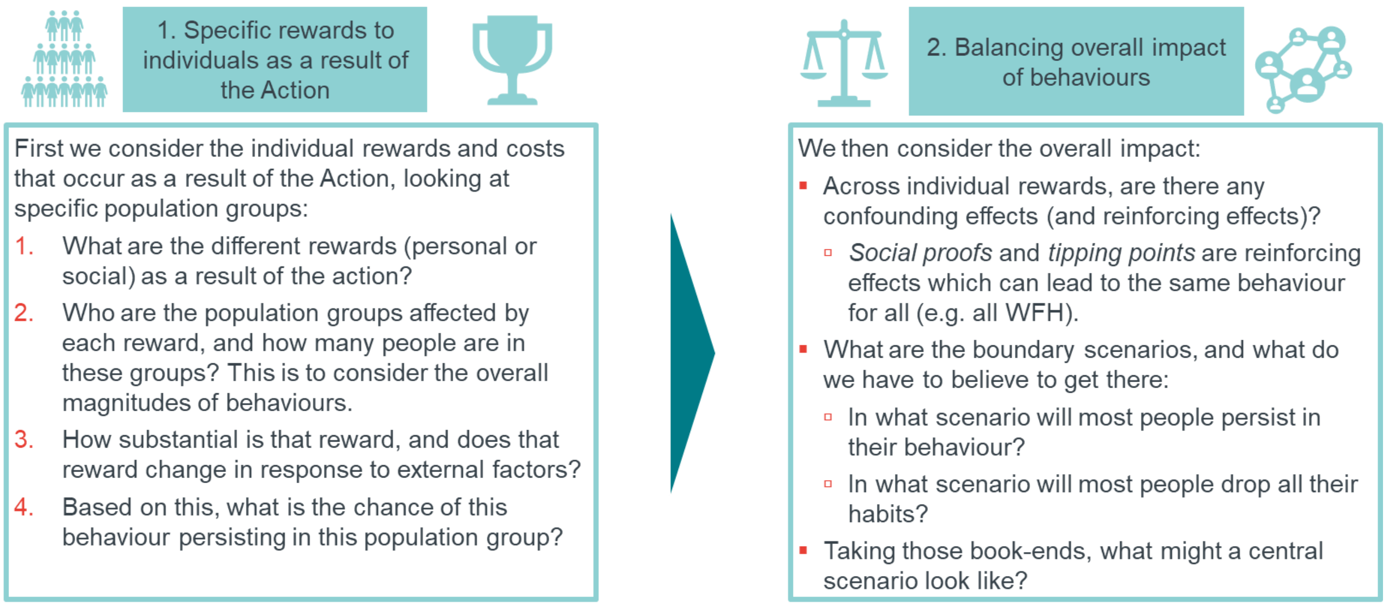
Figure 6 Framework for assessing persistence



Source: Frontier Economics

We apply this framework to the individual behaviours and impacts and the overall impact at GB level (Figure 7). For the overall impact, we also consider the likely scale of the behaviour changes. The actual amount of persistence that continues for each behaviour change will depend on the general level of persistence within the population. Given the uncertainty around future persistence, we consider low, medium and high persistence scenarios for each hypothesis.

Figure 7 Individual behaviours and overall impacts



Source: Frontier Economics

When applying this framework, we are triangulating between a range of evidence sources with varying statistical robustness, relevance and cognitive reliability. To indicate the quality of the evidence, we use the framework set out in Table 3.

Table 3 Evidence quality for triangulation

Statistical robustness	Relevance	Cognitive approach
Quantitative research which covers a large and representative sample, and which should allow reasonable inferences to be made of the population it is sampled from.	The evidence relates exactly or very closely to the information that was required	Research is reporting on **actual events**, or has been has been designed to allow individuals to respond based on their actual experience
Research which is not representative or has a small sample size. This includes **qualitative research** which is not designed to determine how prevalent a view is within the population, but can help generate hypotheses and help understanding of perspectives and drivers	The evidence does not exact exactly to the information that was required and therefore it is a proxy	Research is partly focussed on **hypothetical scenarios**. There may be some uncertainty about respondents ability to give an answer which would reflect their real behaviour.

Source: Frontier Economics

Example: Working from home

To demonstrate our methodology, we now apply the framework to working from home population group described in Figure 5 above. The complete application of the methodology is set out in Annex 1 – Customer behaviour framework for habit formation.

- **Trigger:** There are two potential triggers:
  - For the more risk averse, the initial spread of the virus triggered (proactive) working from home.

- For the less risk averse, government imposed working from home guidelines that triggered working from home.

- **Action:** This trigger was associated with 40-48% of the working population working from home (full time or hybrid). Between May 2020 and May 2021, the percentage of people working from home (full time or hybrid) varied between 27% and 49%, and fluctuated over the period[12].
- **Reward:** Both rewards and costs are relevant to determining the net reward.
  - **Social reward** is the reward to others of working at home, such as reducing the chance of spreading COVID-19 when commuting or in the office.
  - **Personal reward** is the reward to oneself from working at home, such as the ease and convenience of not commuting.
  - **Personal cost** is the cost of working from home, such as loneliness.
- **Ease of persistence:** There are two ways this repeated action has primed continued working from home:
  - Individuals have set up their homes to become workspaces (e.g. creating workspaces at home); and
  - Organisations have learned how to manage remote working (e.g. businesses have adapted to circumstances)

Based on this application of the habit formation framework, Figure 8 sets out our assessment of the impact of increased working from home on an individual basis.

Figure 8 Increased working from home: individual impacts

Rewards from the action	Population group	Size of reward	Chance of persistence
Reducing the chance of spreading COVID when commuting or in the office.	18% of adults felt "uncomfortable" or "very uncomfortable" about leaving their homes as of July 2021. **[1]**	Large reward of staying home before vaccination programme, reducing over time, but some ongoing reward to do so when ill.	Low: Depends on external factors, such as company rules / guidelines on WFH / working in the office.
Saving money	29% of the working population are low income workers who will also WFH [hybrid (26%) or full time (3%)] **[2]**	Greater reward for lower income workers (noting that lower income workers are less likely to WFH)	High: More likely to continue WFH multiple days per week.
Convenience of not commuting (*'opportunity cost'*)	Parents WFH make up 25% of the working population. **[3]**	Greater reward for working parents and workers who have a long commute.	High: More likely to continue WFH multiple days per week.
Reducing chance of catching COVID	1.6% of the working population are WFH to avoid COVID, or because they're clinically vulnerable. **[4]**	Greater reward for those who are, or live with people who are, risk averse or vulnerable	Medium: Different risks for the vulnerable and risk-averse compared to others.

- Survey respondents who said they "*don't have a safe way to travel to work*" or are "*shielding at home due to being clinically vulnerable*".
- Given that we have not found data on the number of employees avoiding their workplace because they are worried about catching COVID, we think that this percentage is at the low end of the plausible range.

High chance of persistence
Medium chance of persistence
Low chance of persistence

Source: Frontier Economics

---

[12] *ONS, Business and individual attitudes towards the future of homeworking, UK: April to May 2021*

Annex 1 – Customer behaviour framework for habit formation sets out our application of this analysis to:

- furloughed workers staying at home;
- elderly and long term unwell shielding at home;
- school-age children studying at home; and
- unemployed remaining at home.

In each case, we examine the impact of the pandemic on electricity consumption, and assess the likelihood of behaviours that drove the impact persisting over the course of ED2.

2.4 Summary of hypotheses

Table 4 sets out a summary of the hypotheses by customer type, generated based on this analysis.

At an overall level, we can see that the pandemic may have resulted in the following changes in electricity consumption from domestic customers:

- an increase in overall consumption;
- changes to the profile of consumption, including:
  - large increases in daytime consumption;
  - small increases in evening consumption;
  - morning consumption shifting to later in the morning; and
- some persistence of these changes over time, but a reversion back to usual behaviours for many customers.

We then go on to test these in Section 3.

Table 4 Hypotheses summary

Group	Potential impact on electricity consumption	Expected likelihood of persistence
Office workers working from home	■ Increase in daytime electricity consumption. ■ Morning electricity consumption shifting later in the morning as commuters save time from commuting and start their days later. The impact on the evening peak depends on the working hours of those who work from home. Home consumption may also be higher in the evening to the	High

Group	Potential impact on electricity consumption	Expected likelihood of persistence
	extent that these workers reduce their normal activities outside of the home.	
Furloughed workers spending more time at home	▪ Increase in daytime electricity consumption. ▪ Overall impact on evening consumption likely to be low, as some workers would have been home then anyway. However, home consumption may also be higher in the evening to the extent that these workers reduce their normal activities outside of the home.	Low
Elderly and long term unwell	▪ Smaller change in daytime consumption than other groups, as may spend more time at home anyway.	Medium
School-age children	▪ Increase in day time consumption ▪ Overall impact on evening consumption likely to be low, as they would tend to be at home anyway.	Low
Unemployed workers	▪ Increase in day time consumption, but this may be lower than for furloughed workers and school age children, due to lower electricity consumption generally for this group.	Medium
Increase in the number of young people moving back home	▪ Increase in consumption in some households, reduction in others.	Low
Increase in the number of non-UK nationals leaving the UK	▪ Overall reduction in electricity consumption	Medium

Source: Frontier Economics

3 Modelling the impact of the pandemic on consumption

In the previous section we set out our hypotheses for how customer behaviour, and therefore consumption, might have changed in response to the pandemic. In this section, we analyse customer-level smart meter data to test these hypotheses.[13] We consider the following:

- consumption changes, relative to pre-pandemic consumption, at different periods of the day, particularly the evening peak (which allows us to understand the impact on network capacity);
- consumption changes, relative to pre-pandemic consumption, for different sub-groups of customers, given different customer groups have different behaviour profiles, and particularly thinking about vulnerable customer impacts; and
- consumption changes over the course of the pandemic, considering the impact is likely to be varied over time according to stages of lockdowns and government rules, which might give us insight into the variability and persistence of these effects.

The structure of this section is as follows:

- We first describe our methodology at a high level. More detail is provided on the modelling approach in Annex 2 - Methodology for estimating the counterfactual.
- We then investigate the changes in overall domestic electricity consumption that may be attributed to the pandemic, both month-by-month and for specific sub-daily periods.
- We then analyse these changes in consumption by customer subgroup. This allows us to understand whether changes vary by subgroup, and to test the hypotheses relating to specific groups. We then use clustering techniques to group similar customer types together, to understand how differences in consumption changes are distributed across the sample of customers. Finally, we use regression analysis to test for relationships between consumption changes and customer characteristics.
- Finally, we use this analysis to understand the potential impact over the short to medium term of the pandemic on:
  - NGED's network capacity; and
  - the bills paid by consumers (especially more vulnerable consumers).

3.1 Methodology

In this section we describe the data, the modelling approach, and the steps used to estimate the impact of the pandemic on electricity consumption. We then describe the steps taken to develop a 'counterfactual'

[13] This section of the report was produced by both Frontier Economics ("Frontier") and the Smart Energy Research Lab (https://serl.ac.uk/) at University College London ("UCL"). Specifically, UCL worked on the counterfactual modelling and impact on customer bills; while Frontier worked on changes in consumption by customer groups and the impact on the network. Both parties worked collaboratively on the overall changes in consumption.

consumption profile (i.e. an estimate of what consumption could have been in the absence of the pandemic) for each household in the sample. Finally we discuss any caveats of the analysis. More detail on this methodology is in Annex 2 - Methodology for estimating the counterfactual.

3.1.1 Smart meter data

The purpose of this report is to understand the persistence of electricity consumption demand patterns developed during the pandemic in the future. Given that, we have a number of data requirements.

First, we require **household-level data** to allow us to explore differences in customer behaviour in response to the pandemic. Our hypotheses suggest that there may be significant diversity in how households changed their behaviour during the pandemic (see Table 4 in Section 2.4 above). To test these hypotheses requires data at a household level. Data that is aggregated across households (e.g. network substation data) does not have sufficient granularity.[14]

We also require **supplementary datasets** alongside the consumption data, to allow us to match relevant household characteristics to consumption profiles. This is to enable us to identify which customers tended to change their consumption over the course of the pandemic, and by how much. Ideally, we would want data that matches to the key characteristics driving behavioural change in our hypotheses set out in Section 2.4. This allows us to investigate how the pandemic may have affected the electricity consumption of specific customer types, for example retired households, or households with children.

We require the household-level consumption data to be a time series, extending back at least a year prior to the start of the pandemic in early 2020 and, ideally, extending to recent periods. Data from before the pandemic is required to allow us to construct a baseline or counterfactual consumption profile for each household. To observe changes we need to track the same households from the period extending from before the pandemic until the most recent period.

We require data that is **representative of NGED's network area**. Since the purpose of this analysis to understand how the pandemic may affect NGED's investment needs, it is important that we focus on changes to household consumption for households living in this area.

To meet these needs, we have used individual-level smart meter data from the Smart Energy Research Lab (SERL)[15]. SERL is a consortium of universities led by UCL. It undertakes research with smart meter data, and provides a secure, consistent and trusted channel for researcher to access high-resolution energy data. This data includes household-level half-hourly electricity consumption data that can be linked to household characteristics, going back to March 2019. The broader SERL data is representative of the general population. We are using a sub-sample of 586 households for which data is available over three years, predominately focused in the south of England. While this means that the aggregate impacts we describe (for example on the network) are not necessarily representative of GB as a whole, the subset we have used as the sample is likely to be more representative of NGED's license areas, which are towards the south of the country. However, our analysis is still relevant at GB level, as more generally, it helps us understand how

[14] If the mixture of households connected to each substation is known then, in principle, regression analysis could be used to tease out the impact of household types. However it would be much harder to identify specific patterns of behaviour without household-level data.

[15] https://serl.ac.uk/

different changes in consumption may correlate to household characteristics. This information is transferable to other areas.

3.1.2 Modelling the counterfactual

The SERL dataset gives us the actual electricity consumption of each household. However, to determine the impact of COVID-19, we need to estimate the counterfactual consumption – what consumption could have been absent the impact of the pandemic. The impact of the pandemic is then estimated as the difference between the counterfactual and actual consumption.

A simple way of estimating the counterfactual would be to look at consumption during a year unaffected by the pandemic (e.g. 2019). However electricity consumption is affected by other factors which vary over time – notably the weather, including temperature, sunshine hours and precipitation levels. Such a simple analysis would therefore risk attributing impacts to COVID-19 which were caused by changes in weather.

We have therefore constructed a model which, based on data from prior to the pandemic (the model "training period"), projects what consumption could have been in 2020 and 2021, given factors such as temperature, precipitation, solar radiation, and daily and seasonal trends.

When calculating the difference between the actual and counterfactual consumption patterns, we consider observations by sub-daily periods for weekdays and weekends in a given month. For example, if we care about sub-period A[16], on weekdays, in April 2020, for a given household, we average all data for sub-period A on weekdays in April, rather than looking at one specific data point. This gives one consumption value for a specific household in April 2020 for sub-daily period A on weekdays. We do this in order to remove some of the error that comes from idiosyncratic usage (for example, if the washing machine or other high-consumption appliance is in use on any particular day).

The raw consumption data is measured is consumption units (Wh) and is compiled half-hourly. We then aggregate the half-hourly data into sub-daily periods, so we sum the consumption in each half-hour for each sub-daily period. We then divide our total sub-daily period consumption by the total number of hours in the sub-daily period. Our units are therefore the average consumption in an given sub-daily period (Wh per hour).

The modelling of the counterfactual has been carried out using LASSO regressions and a neural network. These are established methodologies for projecting electricity consumption. Further details are provided in Annex 2 - Methodology for estimating the counterfactual.

Once we have the counterfactual, we estimate the impact of the pandemic on household electricity consumption by looking at the difference between observed household electricity consumption and the estimated counterfactual household electricity consumption. This means that if a household's electricity consumption changed for any other reason that is not controlled for in the counterfactual, we assume it to be caused by the pandemic.

[16] 23:00 until 06:00

Modelling caveats

We also want to consider some of the caveats associated with the modelling approach, some of which are discussed further in Annex 2 - Methodology for estimating the counterfactual.

- As with any predictive model, there is likely to be some level of bias and error in the predicted results (i.e. the estimated counterfactual). Model **bias** can be thought of as systematic error that leads to a model not matching the training data set closely. For example, if the model systematically underestimated consumption, it would appear that COVID-19 had had more of an effect than it really had. Model **error** measures the inconsistency and uncertainty in results. Out of the two, model bias is the most important factor for us to consider, given we want to minimise the risk of systematic errors in the results. Overall the model shows low bias and the model error is at a reasonable level. Using a combination of modelling techniques helps to negate any impact on the results. The level of model bias and error is described in greater detail in Annex 2 - Methodology for estimating the counterfactual.

- The model training period for the counterfactual starts in March 2019 and ends in February 2020, before behaviour changes associated with the pandemic set in. In estimating what consumption would have been in 2020 and 2021 in the absence of the pandemic, we have taken into account some factors such as weather, which could drive non-pandemic related changes in consumption. However, our model does not take into account all factors that could drive changes in behaviour over time. For example, our modelling does not take into account household changes, such as employment status, number of household occupants, or household income levels that could have occurred over time in the absence of the pandemic. However, these impacts are likely to be relatively small, given the short period of time in question.

- The sample of data is 586 households. This is likely to be large enough to draw some illustrative conclusions about overall consumption changes during the pandemic. However it becomes more difficult to isolate the impact of the pandemic on sub-groups of customers due to the small sample size in specific groups. For some sub-groups, the sample reduces to less than 100, which is unlikely to be enough observations to draw firm conclusions on the impact of the pandemic. Where this is the case, we note it in our results.

Despite the caveats listed above, we believe this is a good dataset to undertake an analysis of the pandemic impact on household electricity consumption, and can be used to identify differences in household behaviours.

3.2 Overall changes in consumption

This section describes changes in household level consumption during the pandemic.

- We first recap on our hypotheses and the findings from the literature review.

- We then consider aggregate changes in consumption. We consider all periods of all the day for all months in our sample, which is from April 2020 to December 2021.

- Once we have determined when there is a difference in consumption between actual consumption and estimated counterfactual consumption and the size of the difference, we then look at which periods of the day may explain the difference in consumption. Finally, we look at how the difference in consumption for two key periods of the day, the daytime and the evening, vary across the sample.

Our focus in this section is on overall customers. In later sections, we consider how this varies by customer sub-groups. The hypotheses developed in Section 2.4 predict why different groups of people may stay at home due to the pandemic and the associated restrictions.

Taken in the round, the hypotheses for different groups of customers suggests that we may see the following changes in electricity consumption for households at an aggregate level:

At an overall level, we can see that the pandemic may have resulted in the following changes in electricity consumption from domestic customers:

- an increase in overall demand;
- changes to the profile of demand across the day, including:
  - large increases in daytime consumption;
  - small increases in evening consumption;
  - morning consumption shifting to later in the morning; and
- some persistence of these changes over time, but a reversion back to usual behaviours for many customers.

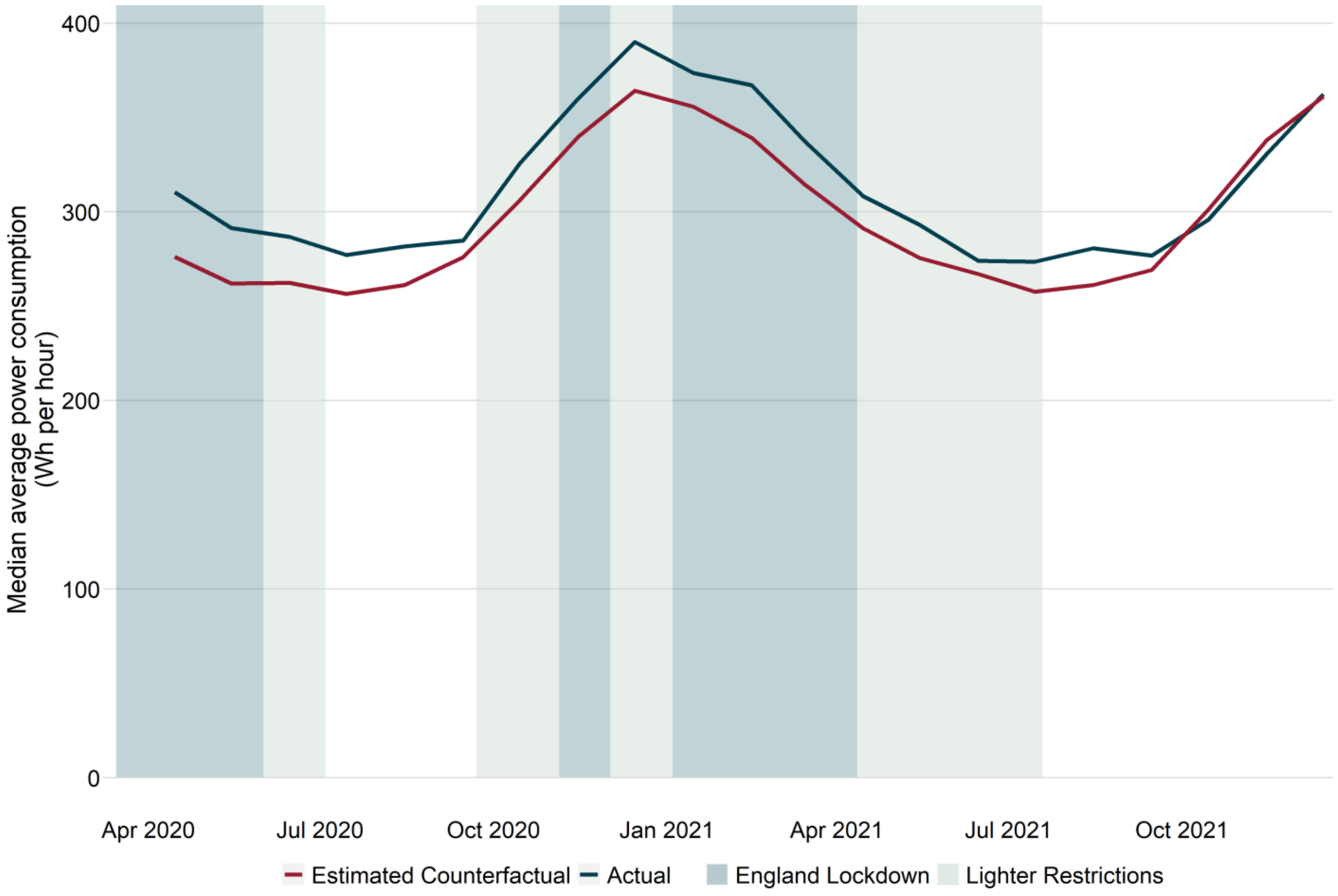
3.2.1 Overall demand

In this section we describe the overall impact we observe on monthly demand, and the implications for our hypotheses.

Descriptive analysis

We first examine the overall change in consumption across all customers in our sample. Figure 9 below shows actual and estimated counterfactual household consumption from April 2020 to December 2021. This shows that actual consumption was consistently higher than estimated counterfactual consumption throughout 2020, and at the beginning of 2021. This is consistent with the research described in Section 2.2.1 showing an increase in overall residential electricity consumption during lockdowns. The difference between actual and estimated counterfactual consumption then decreased during the summer and autumn of 2021.

Figure 9 Actual and estimated counterfactual consumption – April 2020 to December 2021



Source: Frontier Economics / UCL analysis using SERL data

It can be difficult to determine the impact of specific lockdowns from Figure 9 because the difference between the lines can be quite small compared to the overall magnitude of the lines. We therefore show other charts in terms of the *percentage* difference on the estimated counterfactual consumption. This should give greater insights into changes in customer behaviour than the absolute difference. For example, if a certain proportion of households started working from home, we would expect the increase in consumption to be highest during times of the year and day where consumption was already higher.

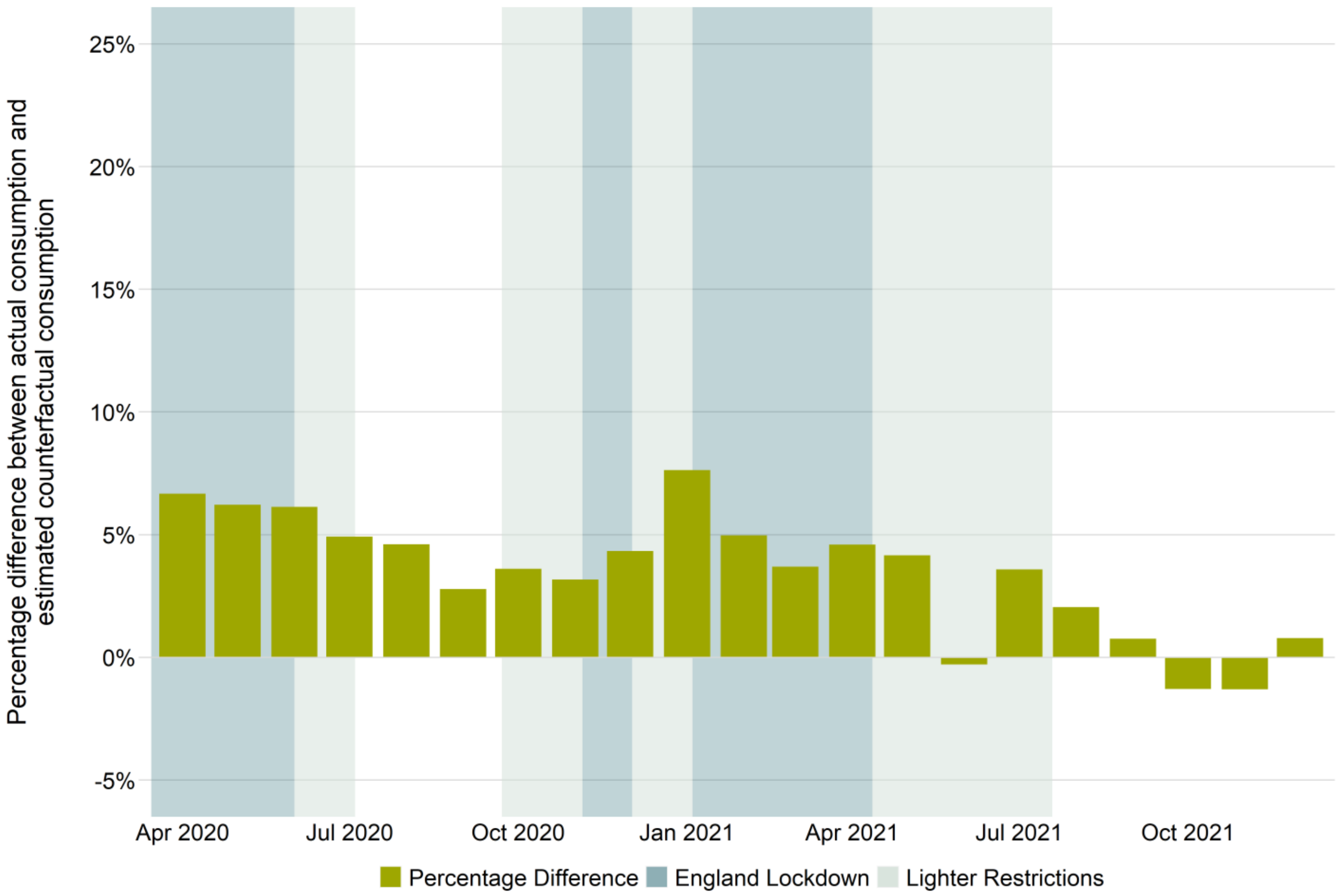
Figure 10 shows that actual consumption was consistently higher than estimated counterfactual consumption over the course of 2020, as we also saw in Figure 9. In other words, household consumption was higher than we estimate that it would have been if the pandemic had not occurred. The percentage difference between actual consumption and the estimated counterfactual was particularly high in April 2020, with actual consumption around 6.5% higher than estimated counterfactual consumption.

Figure 10 shows that the large difference in consumption decreased to some extent into the summer of 2020, but still remained positive, meaning that actual consumption was above estimated counterfactual consumption. The difference between actual consumption and the estimated counterfactual decreased from 6% in June 2020 to approximately 3% in September and October 2020.

The difference between actual consumption and the estimated counterfactual then increased into the winter of 2020 as high virus prevalence resulted in a lockdown during most of November, and a lockdown for the first few months of 2021. The difference in consumption then decreased during the first half of 2021 as restrictions eased, with a slightly negative difference in June 2021. Despite the continued easing of restrictions, with the vast majority of restrictions easing on 19th July, the difference increased considerably into July, reaching approximately 4%. This correlates with a large increase in cases (and thus virus prevalence) at that time.

The difference between actual consumption and the estimated counterfactual then decreased into the autumn of 2021 as the number of cases remained steady until the end of November. The difference in consumption became negative for October and November 2021, which means that estimated counterfactual consumption was higher than actual consumption, with a small positive difference in December 2021.

Figure 10 Percentage difference in monthly consumption – April 2020 to December 2021



Source: Frontier Economics / UCL analysis using SERL data

Implications for our hypotheses

These findings align with our hypothesis that overall household consumption increased during the pandemic. The large percentage difference between actual consumption and the estimated counterfactual consumption in the spring of 2020 is likely caused by the full lockdown. The lockdown led to both adults and school-age children staying at home due to work from home orders and the closure of schools, respectively. In addition, this relatively large difference in consumption may be partly due to workers who are unable to work from home staying at home under the furlough scheme.

The decrease in the difference in consumption during the summer of 2020 may be explained by the re-opening of schools, offices, shops and hospitality which occurred in June and July 2020 as restrictions were eased. This is consistent with our hypotheses around the persistence of some of the change in consumption: in particular, those relating to school-age children and workers who stayed at home under the furlough scheme who could then return to work in industries such as retail and hospitality. Actual consumption remaining above estimated counterfactual consumption as restrictions eased suggests that some groups of people did not fully return to their pre-pandemic behaviour, and some of the changes in behaviour that occurred during the lockdown of spring 2020, such as working from home, continued to some extent into the summer and the beginning of autumn.

In July 2021, the difference between actual and estimated counterfactual consumption increased as cases rose but restrictions eased. This suggests that virus prevalence might be a second key driver of changes in the difference in consumption, with lockdowns and the associated restrictions being the other key driver (see 4.1 and Annex 3 – changes in consumption for charts of the change in consumption for different periods of the day and estimated COVID-19 cases in England).

The small difference in consumption towards the end of 2021 suggests that any changes in behaviour that occurred during lockdowns may have almost or fully reverted at an aggregate level. However, it is still possible that changes in behaviour for specific groups of customers may have persisted, despite the aggregate picture suggesting little persistence. We examine changes in consumption for specific groups of customers in the next section.

3.2.2 Changes to the profile of demand across the day

Having examined how consumption has changed across months, with actual consumption greater than estimated counterfactual consumption for 2020 and the first half of 2021, we now examine how consumption has changed across a typical weekday[17]. This enables us to determine the importance of specific periods of the day in accounting for the difference between actual consumption and the estimated counterfactual.

Descriptive analysis

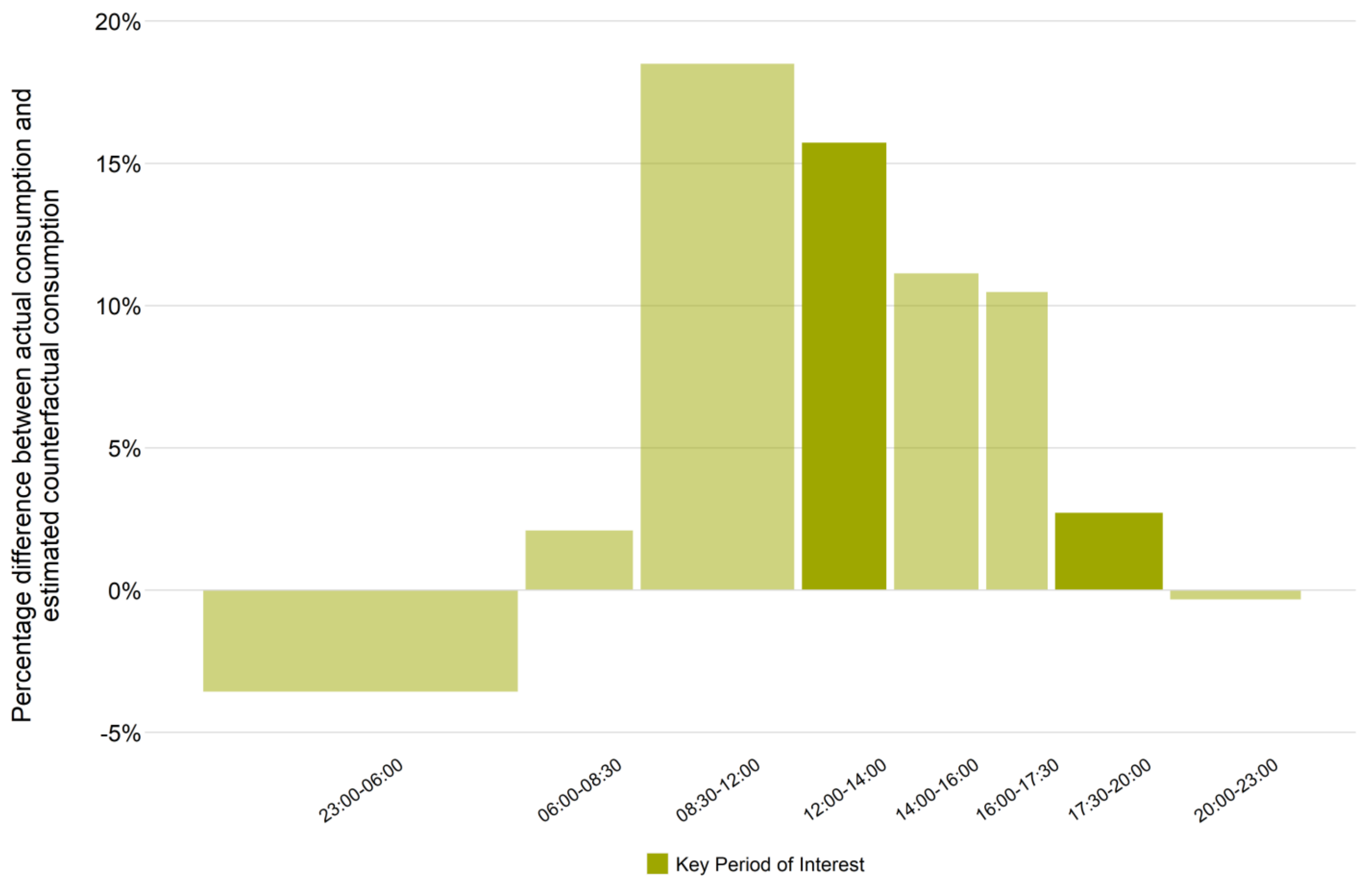
It is important to determine which periods of the day account for the difference in consumption, as peak consumption, which we define as between 17:30 and 20:00, is most important for network operators who need to build their networks to meet this demand.

[17] This is explained in Section 3.1.2

We examine changes in consumption across a typical weekday for April 2020, which is a month characterised by a full lockdown, with schools shut, a work from home order for those who were able to work from home, and the closure of non-essential shops and hospitality leading to some workers staying at home under the furlough scheme. We also examine changes in consumption across a weekday for November 2021, when no lockdowns or restrictions were present.

Figure 11 shows how the difference between actual and estimated counterfactual consumption changed across eight sub-daily periods of a typical weekday in April 2020. This shows that the greatest differences in consumption occurred during the daytime. The largest difference in consumption occurred from 08:30 – 12:00, with actual consumption around 18% higher than estimated counterfactual consumption. The difference in consumption continued to be relatively large for the next three sub-daily periods, which span 12:00 – 17:30, with the difference between actual and estimated counterfactual consumption ranging from 11% - 16%. The difference was considerably lower between 17:30 – 20:00 compared to the daytime differences in consumption, with actual consumption only 2%-3% higher than estimated counterfactual consumption. The difference in consumption then turned very slightly negative between 20:00 – 23:00, which means that estimated counterfactual consumption was greater than actual consumption. We highlight two sub-daily periods we are particularly interested in: between 12:00 – 14:00 (since it represents changes in daytime demand) and 17:30 – 20:00 (since it represents changes in evening peak demand).

Figure 11 Percentage difference in consumption – an average weekday in April 2020



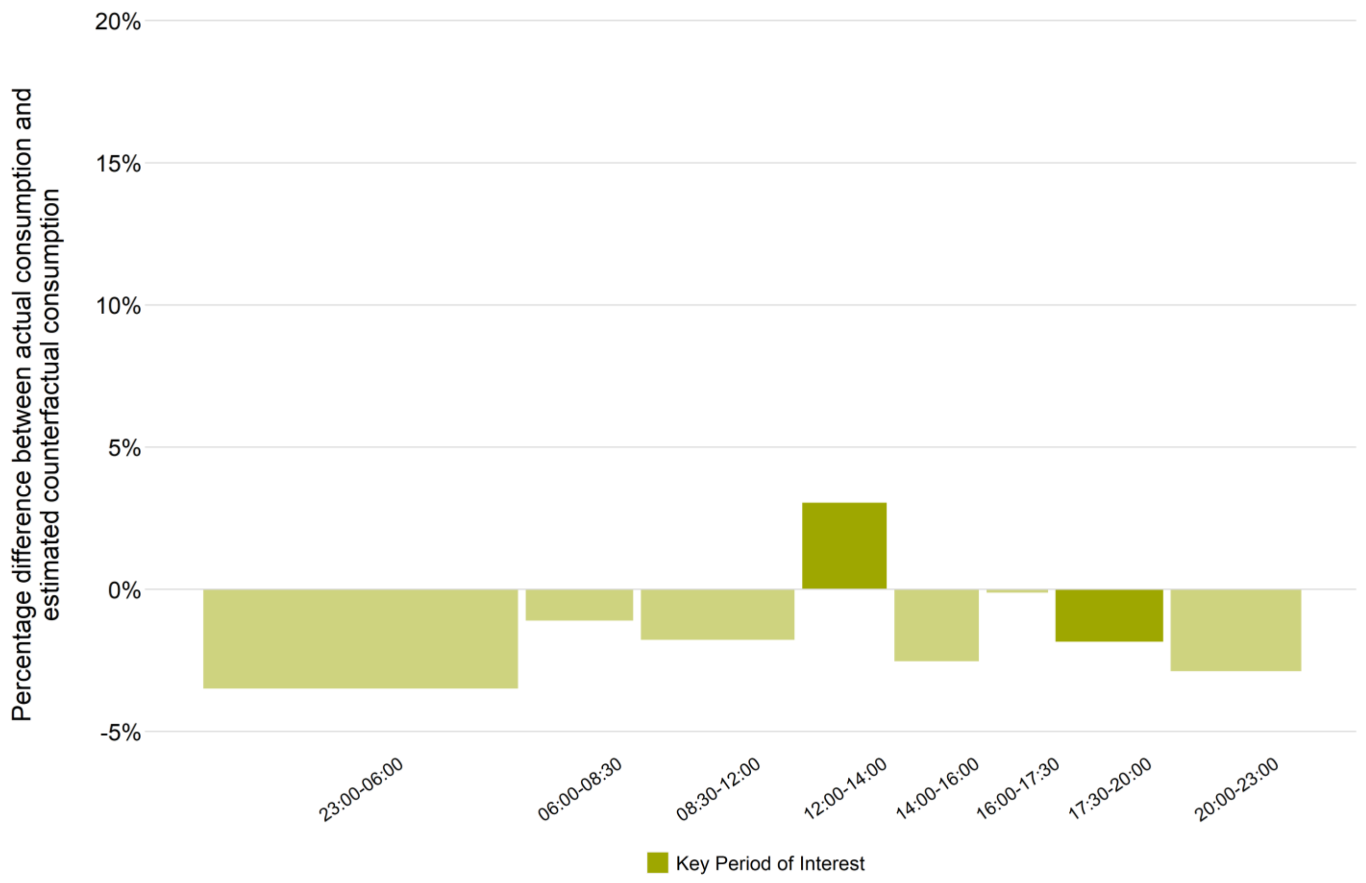
Source: Frontier Economics / UCL analysis using SERL data

Figure 12 shows how the difference between actual and estimated counterfactual consumption changed across eight sub-daily periods of the typical weekday in November 2021. During this month there were no restrictions, and there was limited virus prevalence.

Figure 12 shows that most differences in consumption were small in magnitude, with all but one sub-daily periods showing negative differences in consumption, meaning that estimated counterfactual consumption is greater than actual consumption, and that the pandemic appeared to have driven a reduction in consumption. The only sub-daily period where actual consumption was greater than estimated counterfactual consumption was 12:00 – 14:00, with actual consumption around 3% higher than estimated counterfactual consumption. The difference in consumption was slightly negative from 14:00 – 17:30, and the difference then became more negative, with estimated counterfactual consumption greater than actual consumption by 2% and 3% between 17:30 – 20:00 and 20:00 – 23:00, respectively. The difference in consumption was also negative in the morning. Estimated counterfactual consumption was greater than actual consumption by approximately 1% and 2% between 06:00 – 08:30 and 08:30 – 12:00, respectively.

The changes in consumption across the typical weekday in November 2021 were notably different to the changes across the day in April 2020. Whereas actual consumption was greater than estimated counterfactual consumption for most periods of the day in April 2020, with particularly large differences during the day, estimated counterfactual consumption was greater than actual consumption for most periods of the day in November 2021.

Figure 12 Percentage difference in consumption - an average weekday in November 2021



Source: Frontier Economics / UCL analysis using SERL data

Implications for the hypotheses

These findings align with our overall hypotheses that the pandemic would result in changes to the profile of demand across the day, including: large increases in daytime consumption and small increases in evening consumption. However, it does not look like early morning consumption shifted to later in the morning.

The analysis of changes in consumption across the day for April 2020 and November 2021 suggests that:

- there can be a large difference in daytime consumption when restrictions are in place, such as during lockdowns; however
- there are only small changes in evening (peak) consumption, regardless of whether a lockdown occurs.

Persistence is difficult to assess using just this analysis. The small differences in consumption across the day in November 2021 suggests that any changes in behaviour that occurred during lockdowns may have almost or fully reverted at an aggregate level. As previously noted, it is still possible that changes in behaviour for specific groups of customers may persist, despite the aggregate picture suggesting little persistence of behaviours that may change consumption. We investigate this further in Section 3.3.

The negative difference in consumption across the early morning and afternoon, which means that estimated counterfactual consumption was greater than actual consumption, isn't suggested by any of the hypotheses set out in Section 2.4. There are two possible reasons for the unexpected negative difference in consumption:

- The pandemic may have caused households to change their behaviour in other ways that have not been picked up in our hypothesis generation.
- There may be other factors which mean that some groups of customers reduced their consumption. As discussed in Section 3.1.2 and in Annex 2 - Methodology for estimating the counterfactual, consumption patterns may have changed independently of the pandemic, and of the other factors we take account of in our analysis. This is more likely to occur later in the sample period, in other words, towards the end of 2021, because counterfactual household consumption in towards the end of 2021 is less likely to be comparable to the household consumption patterns in 2019.
- The random errors in the model (described in Section 3.1) may be driving the result. Although, as indicated in Annex 2 - Methodology for estimating the counterfactual, the model appears to have very low bias for November, which could not explain the results reported above.

We now examine how daytime and evening differences in consumption change over time.

3.2.3 Changes in daytime and evening consumption over time

Having examined how overall consumption has changed, and which periods of the day account for these changes in consumption, we now consider how consumption in two key periods of the day, the daytime and the evening, varied over time.

It is important to understand changes in daytime consumption as many of our hypothesised changes in behaviour (e.g. working from home) may affect daytime consumption. The level of consumption across the day can be an important determinant of the size of customers' bills[18], which is especially important for certain groups of customers from a vulnerability perspective. In addition, changes in daytime consumption are a useful indicator regarding whether individual load profiles have changed.

As noted above, consumption during the evening peak is a key driver of network reinforcement requirements.

Descriptive analysis

As discussed at the beginning of Section 3.2, our hypotheses suggest that in the daytime, actual consumption should be considerably larger than the estimated counterfactual in periods of restrictions, such as lockdowns. The extent to which this difference continues in the months following periods of restrictions is likely to be a useful indicator of the level of persistence of some behaviour changes.

Figure 13 shows how daytime consumption, 12:00 – 14:00, changes between April 2020 and December 2021. There is a similar pattern to the aggregate picture, with:

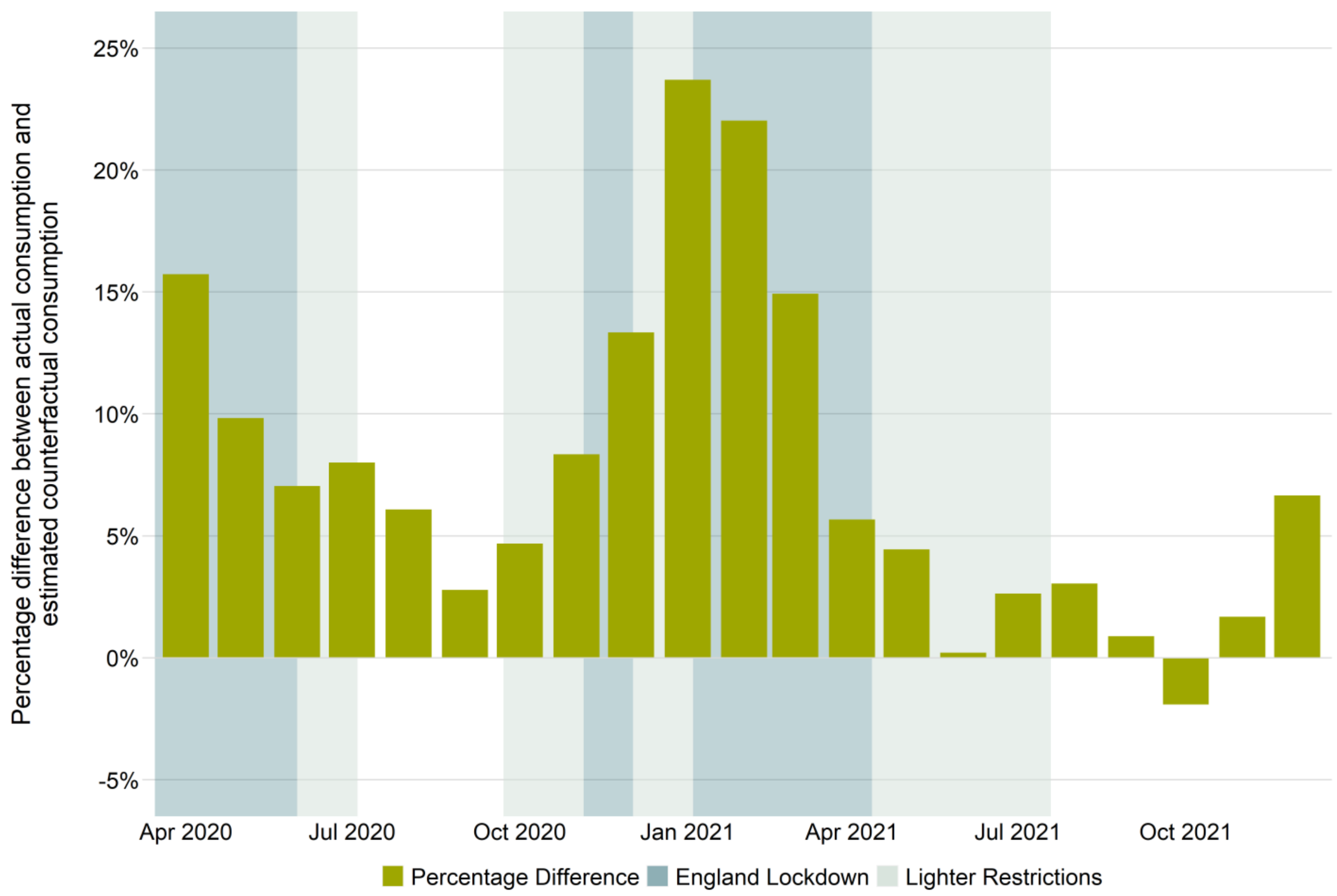
[18] Depending on the tariff choice of individual customers

- Actual consumption greater than estimated counterfactual consumption across 2020 and the beginning of 2021, with a greater difference during the periods of lockdown, such as the spring of 2020 and the beginning of 2021.
- The difference in consumption then decreased in magnitude from the beginning of 2021 to the autumn of 2021.

In general, the changes in consumption during this period seem to exhibit a much greater correlation with periods of lockdown than the aggregate figures shown in Figure 10 suggesting that, as we expect, much of the overall change in energy use has been driven by daytime consumption.

There were some months where the changes in consumption for daytime consumption differed from the aggregate picture. For example, there was a positive difference for daytime consumption in November 2021, while there was a negative difference for this month at the aggregate level.

Figure 13 Percentage difference in daytime consumption (12:00 - 14:00) - April 2020 to December 2021



Source: Frontier Economics / UCL analysis using SERL data

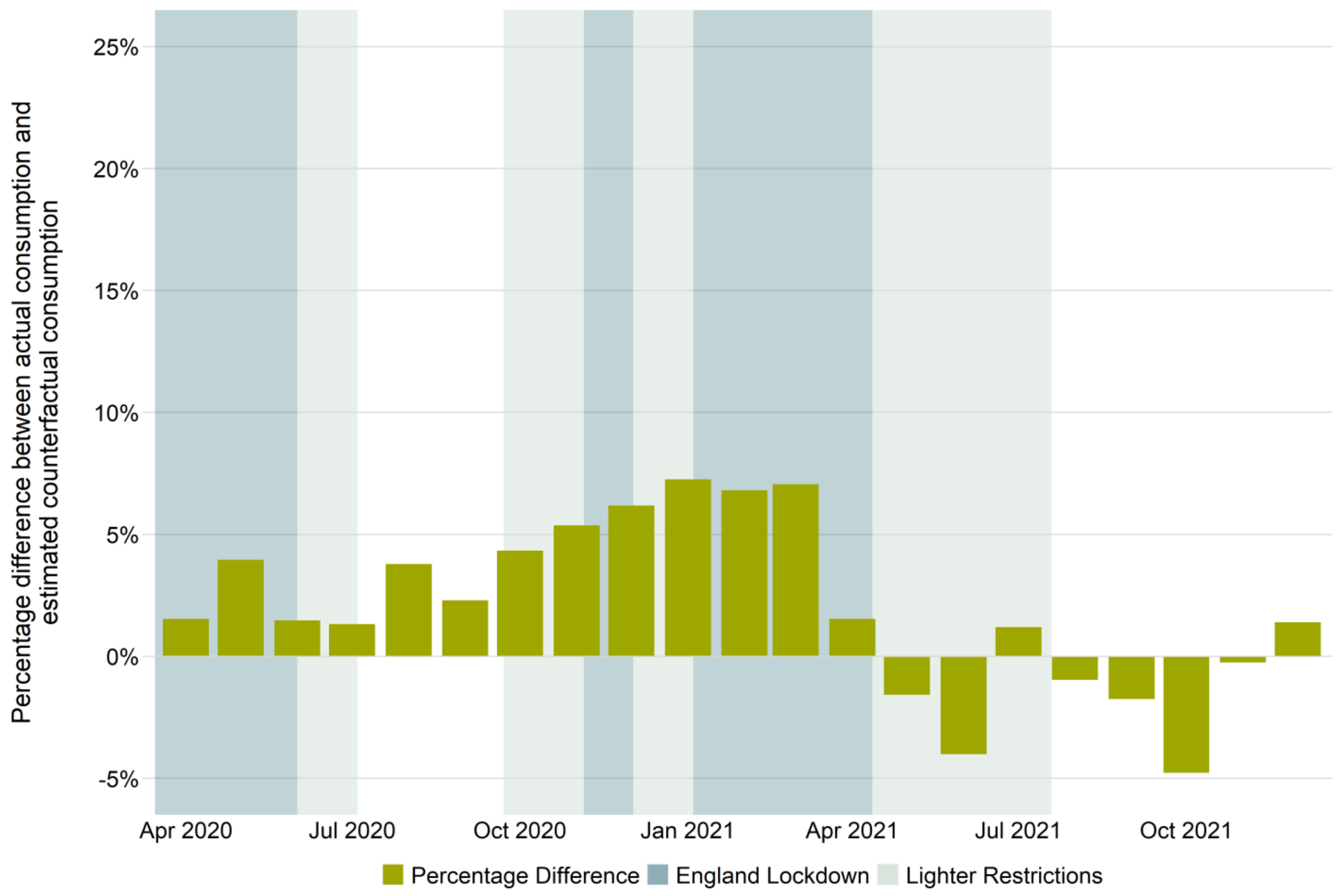
We now examine how differences in consumption in the evening vary over time.

As discussed at the beginning of the ‘overall changes in demand’ sub-section in Section 2, our hypotheses suggest that in the evening, lockdowns may have driven a smaller change in electricity consumption than during the day. While more people would stay in in the evenings during a lockdown, this effect is likely to be smaller than in the day time. For example, most workers would have gone to their workplace on most weekdays in the absence of the lockdown, but would only have gone out on some evenings.

Figure 14 shows how evening consumption, 17:30 – 20:00, changed between April 2020 and December 2021. The difference in consumption was small and positive across 2020, which means that actual consumption was greater than estimated counterfactual consumption. Excluding July 2020, the difference in consumption was of a similar magnitude in spring 2020 and the summer of 2020, despite a lockdown occurring in the spring, and fewer restrictions occurring in the summer. For example, actual consumption was greater than estimated counterfactual consumption by approximately 1.5% in April 2020, and the corresponding value is approximately 3.5% in August 2020. The difference then increased towards the end of 2020 as cases rose and more restrictions were imposed.

For 2021, actual consumption was greater than estimated counterfactual consumption at the beginning of the year when a lockdown occurred, however the difference effectively disappeared by April 2021, and was negative from June 2021 until November 2021.

Figure 14 Percentage difference in evening consumption (17:30 - 20:00) - April 2020 to December 2021



Source: Frontier Economics / UCL analysis using SERL data

Implications for the hypotheses

For daytime consumption, the large differences between actual consumption and estimated counterfactual consumption are consistent with our hypotheses. The smaller difference during the summer of 2020 suggests that while some behaviours during the first lockdown reverted, some behaviours persisted. On the other hand, the small differences in consumption during the summer of 2021 suggests that fewer changes in behaviour persisted following the lockdown at the beginning of 2021. There is a sudden increase in the difference in December 2021, which is a month where cases rose rapidly, suggesting that daytime consumption is sensitive to changes in the number of cases.

For evening consumption, the small differences between actual consumption and estimated counterfactual consumption are consistent with our hypotheses. The consistent level of the consumption difference across 2020 suggests that behaviours that occurred during the spring lockdown, such as people who work long hours working from home, persisted for the whole year. In addition, the estimated counterfactual consumption was greater than actual consumption from the summer of 2021 until November 2021. This suggests that any pandemic behaviours which resulted in the positive difference in consumption during 2020 and the beginning of 2021 were transitory and did not continue from the summer of 2021.

3.2.4 Conclusions

Our analysis of overall changes in consumption suggests the following.

- The pandemic drove an increase in home electricity consumption, but these changes did not persist at an aggregate level when lockdowns were removed. Actual consumption was greater than the estimated counterfactual across 2020 and the beginning of 2021, particularly in periods of restrictions, however this difference dissipated during the second half of 2021.

- During periods of restrictions, such as the first lockdown, the largest differences in consumption occurred during the daytime, with small differences in the early morning and evening. During periods of no restrictions, such as November 2021, differences in consumption were small and sometimes negative (i.e. the pandemic drove a decrease in home consumption), which is in contrast to sub-daily periods during periods of restrictions, which were typically positive.

- These changes vary with changes in lockdown restrictions. The difference in daytime consumption was typically large during periods of restrictions, and was smaller (yet still positive) when restrictions eased, suggesting limited persistence. The difference in evening consumption is typically smaller and more independent of the level of restrictions. The large difference in daytime consumption, smaller difference in evening consumption, and largest impact during periods of restrictions is consistent with our hypotheses set out in Section 2.4, and discussed in Section 3.2.

Our analysis suggests that by the end of 2021, any persistence of behaviour changes that occurred during periods of restrictions is limited. However, as previously discussed, it is still possible that changes in behaviour for specific groups of customers may persist, despite the aggregate picture suggesting little persistence. We examine changes in consumption for specific groups of customers in Section 4, to determine whether behaviour changes have persisted for specific groups.

3.3 Changes in consumption by customer groups

In this section, we present the results of analysis designed to further investigate the shifts in consumption described above at an aggregate level, so we can understand which types of customers changed consumption and how.

- First, we carry out a similar analysis to the aggregate work described above, but splitting the sample of customers by factors which are observable within the SERL dataset, such as the presence of children within the household. Within this analysis we undertake regression analysis. This helps to unpick the impact of multiple factors (such as the presence of children, and the presence of adults of working age) which are likely to be correlated and identify which characteristics may be driving the change in consumption.

- Second, we carry out clustering analysis that groups customers according to the type of response they had to the pandemic, rather than by observable demographic factors. This can help us understand the extent of heterogeneity within clusters (or groups) of customers with a similar response, as well as determining “typical” patterns of behaviour. For example, we test whether households who increased consumption in the first lockdown also increased consumption in the second.

The purpose of this analysis is to help us better understand the drivers of the overall changes in consumption, and ultimately feed into the persistence assessment described later in section 4.

3.3.1 Changes in consumption by customer characteristics

In this section we outline the different customer characteristics and assess changes in consumption, considering the different groups outlined in our hypotheses (outlined in the box below). We use linear regression analysis[19] to understand the relationships between changes in consumption and observable characteristics. We follow the same pattern as the aggregate analysis, looking at:

- the overall load profile at start and end of the period;
- changes in peak demand over time; and
- changes in daytime demand over time.

We use the regression analysis to assess how these outcomes vary by three categories of customer characteristics:

- Ages of household:
  - Retired households, measured as households made up entirely of retirees.
  - Households with children, defined by having at least one occupant aged 15 or below.
- Employment status and home-working:
  - Households where at least one occupant has a degree.
  - Using answer from COVID-19 survey on homeworking, where at least one adult responded saying they work from home.
- Income:
  - Perceived financial security, measured by a survey question asking if the household was financially struggling in some way.
  - Whether the household is located in a region with a high index of multiple deprivation[20].

[19] Regression analysis is a statistical method which estimates the causal relationship between a variable of interest and other variables which may impact the variable of interest.

[20] Defined as a Lower Super Output Area in the bottom quintile on the Index of Multiple Deprivation

We expect to find the following features in household behaviour

- Households that tend to stay at home, who previously did not stay at home, are more likely to significantly increase their daytime demand compared to groups that do not stay at home. Specifically, we identify a subset of groups that might spend more time at home:
  - those households with people working from home (to measure this, we use proxies based on survey data and assumptions on the types of people that work in an office)[21];
  - furloughed workers spending more time at home;
  - Elderly people and those with long term sickness will spend more time at home, and potentially shielding;
  - households with school-age children (who may be home-schooled); and
  - unemployed workers.
- We also identify a subset of groups that tend to change residence, including a subset of young people and non-UK nationals. However, it was difficult to identify these groups in the data.
- Within the affected groups identified, the largest changes in behaviour are likely to coincide with particular events associated with the pandemic, such as when: furlough was active; schools were closed; lockdowns were enforced; and there was a high virus prevalence.
- We expect different levels of persistence by group, for example:
  - households with children may not persist in their behaviour, given schools reopened.
  - retired households may show different levels of persistence, depending on how they perceive risk – but those most vulnerable may persist in staying at home.
  - different levels of persistence are expected for different kinds of home-workers, for example we might expect on average that:
    - younger households are more likely to return to the office[22].
    - those on higher incomes are more likely to remain at home[23], since they may be more likely to have more comfortable working conditions.

[21] We might expect those working in office jobs to be working from home, while office jobs are typically associated with having a degree.

[22] Workers aged 16-29 had a more negative net sentiment towards homeworking than workers aged 30 to 49 and 50 to 69 with respect to distractions, thinking of news ideas, job opportunities, and ease of working with others. In addition, workers aged 16-29 had a less positive net sentiment towards homeworking than workers aged 30 to 49 and 50 to 69 with respect to work life balance and speed of completing work. (ONS, Business and individual attitudes towards the future of homeworking, UK: April to May 2021).

The following sub-sections assess the change in consumption for each category of customer characteristic, estimating the impact of the pandemic on changes in daytime and evening consumption over the whole period (using linear regression analysis), and the implications for our hypotheses.

To assess the change in consumption for each category of customer characteristic, we use the approximate median household's change in consumption. We also use approximations for medians throughout this analysis by taking the average of at least 10 households with a value very close to the median, so that the value for an individual household is not revealed. It is worth noting throughout this analysis that there is significant variation across different characteristics, therefore we cannot generalise conclusions to all households with a specific characteristic. However it illustrates the broad differences between typical households.

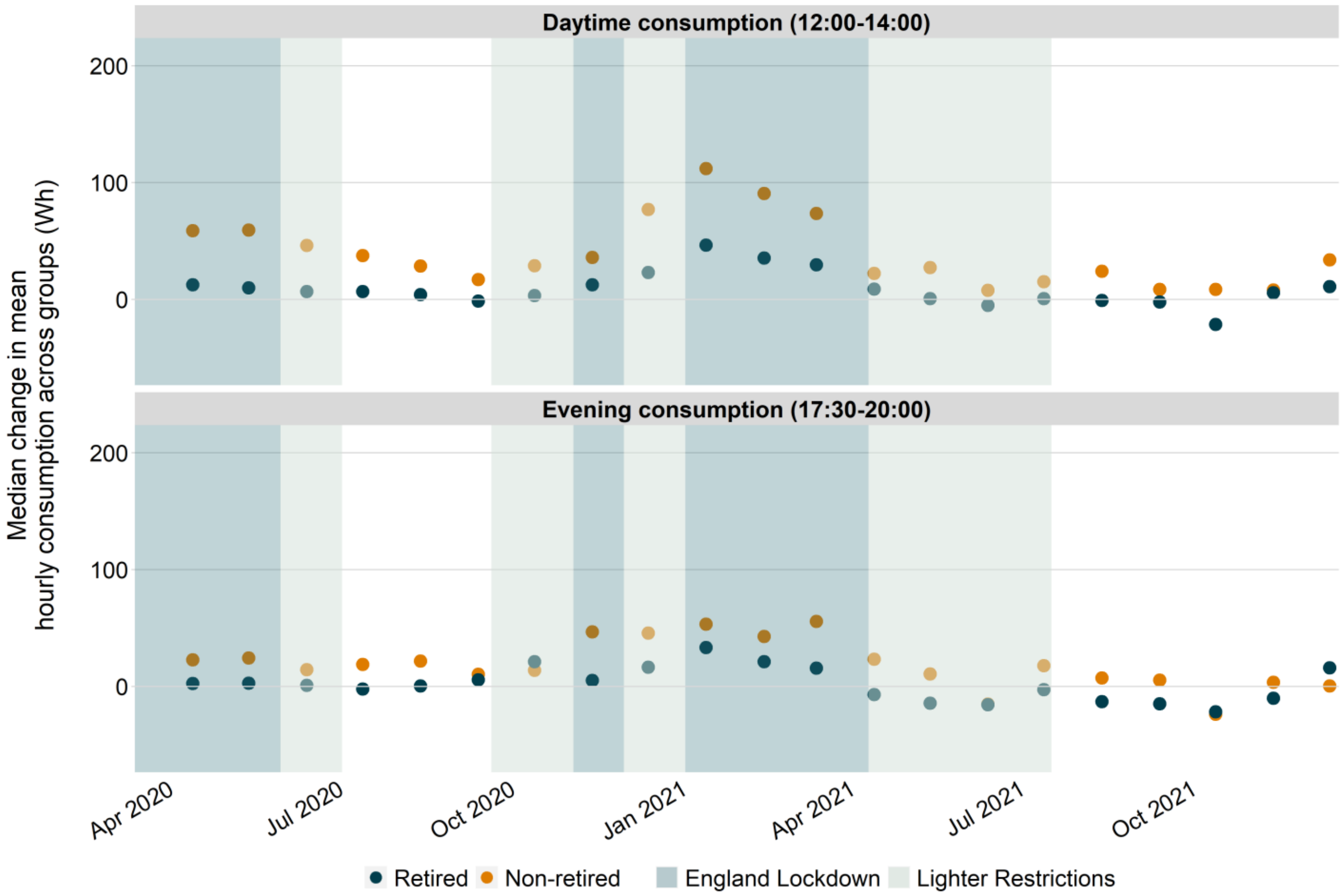
Ages of household

Retired households

The first group of customers we consider are retired households. Figure 15 shows how daytime and evening consumption change over time for retired households compared to non-retired households.

[23] 65% of workers with incomes below £10,000 expected to either hybrid work or not return to their usual place of work. On the other hand, 94% of workers with incomes above £50,000 expected to either hybrid work or not return to their usual place of work. (ONS, Business and individual attitudes towards the future of homeworking, UK: April to May 2021).

Figure 15 Change in daytime and evening consumption – split by households which have retired and non-retired members



Source: Frontier Economics / UCL analysis using SERL data

The change in daytime consumption followed a similar shape for both types of households over the sample, which is a similar shape to overall daytime consumption changes shown in Figure 13. Retired households (in blue) tended to show lower changes in consumption across the whole period; whilst other households showed larger changes in consumption during spring 2020 and winter 20/21[24]. By the end of 2021, there were some remaining differences across groups, although these differences were smaller compared to the lockdown period.

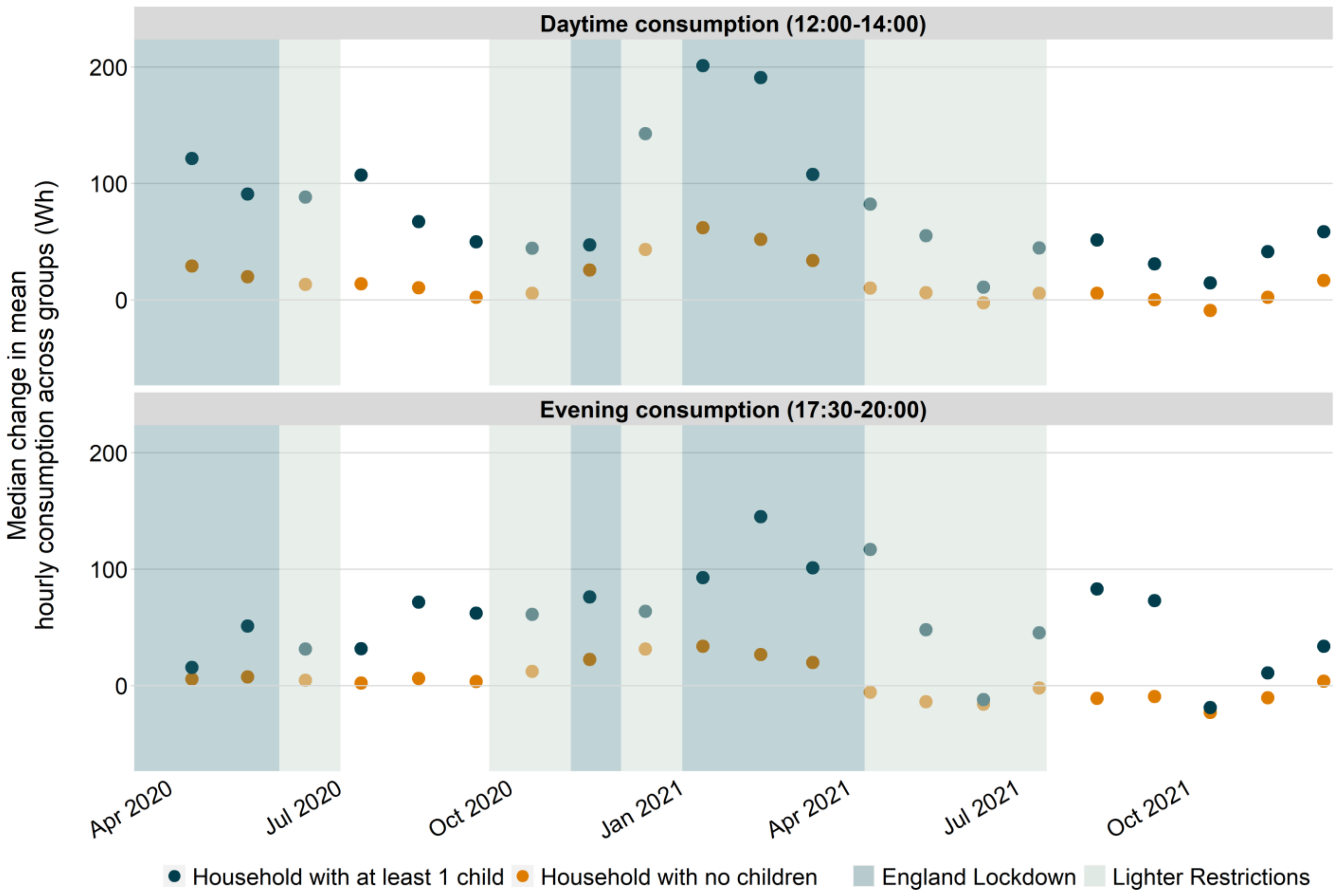
The change in evening consumption followed a similar overall pattern to the changes in daytime consumption. The changes in consumption tended to be smaller across the months for both retired and all other households, although the differential between the two groups was similar over the entire period. Later in the sample there were negative changes in evening consumption for some months for retired households, suggesting that consumption after the pandemic was lower than prior to the pandemic.

[24] The average household counterfactual consumption by customer group, as seen in Annex 3, shows that retired households had similar levels of consumption to other customer groups. This suggests that the driver of lower consumption changes is due to limited behaviour changes, rather than a function of lower consumption.

Household with children

We also consider the same analysis for households with children in Figure 16.

Figure 16 Change in daytime and evening consumption – split by households with and without children



Source: Frontier Economics / UCL analysis using SERL data

The change in daytime consumption tended to be larger for the households identified in this chart compared to the previous chart. There were larger differences between the types of households during early 2020 and winter 2020/21 (corresponding to the lockdown periods), but there was some convergence during summer 2020 and the summer and the beginning of autumn of 2021 (in other words, outside of the lockdown periods).

The change in evening consumption was different to changes in daytime consumption. For example, households with children had larger increases in evening consumption in the spring 2020 lockdown and the summer of 2021 compared to the change in daytime consumption. Whilst households with children had large increases in evening consumption during winter 20/21, they had notably larger increases in daytime consumption during the same period. Households with children had large changes in evening consumption during summer 2021 which do not continue into the autumn. This may reflect households with children temporarily changing typical behaviour during summer, such as being on school holidays but not being able to go away.

Implications for our hypotheses

The results above tend to match our hypotheses. For retired households, the impact of the pandemic on their behaviour was small since these households were more likely to stay at home in the pre-pandemic period. Therefore, any changes in consumption due to the pandemic across the sample of retired households are likely to be small in comparison to other household types.

Households with children tended to use more electricity in the daytime, particularly during the lockdown periods, which reflects children being home schooled. This is consistent with the results of our regression analysis.[25]

One additional effect not in our hypotheses was the large change in evening consumption for households with children. This change in consumption was larger than expected and tended to last as long as changes in daytime consumption. However, changes in consumption in the non-lockdown period for households with children remained higher than changes in consumption for childless households. It is difficult to disentangle this effect: whether it is due to these households having children, or some other characteristic that may be correlated.

Employment status and home-working

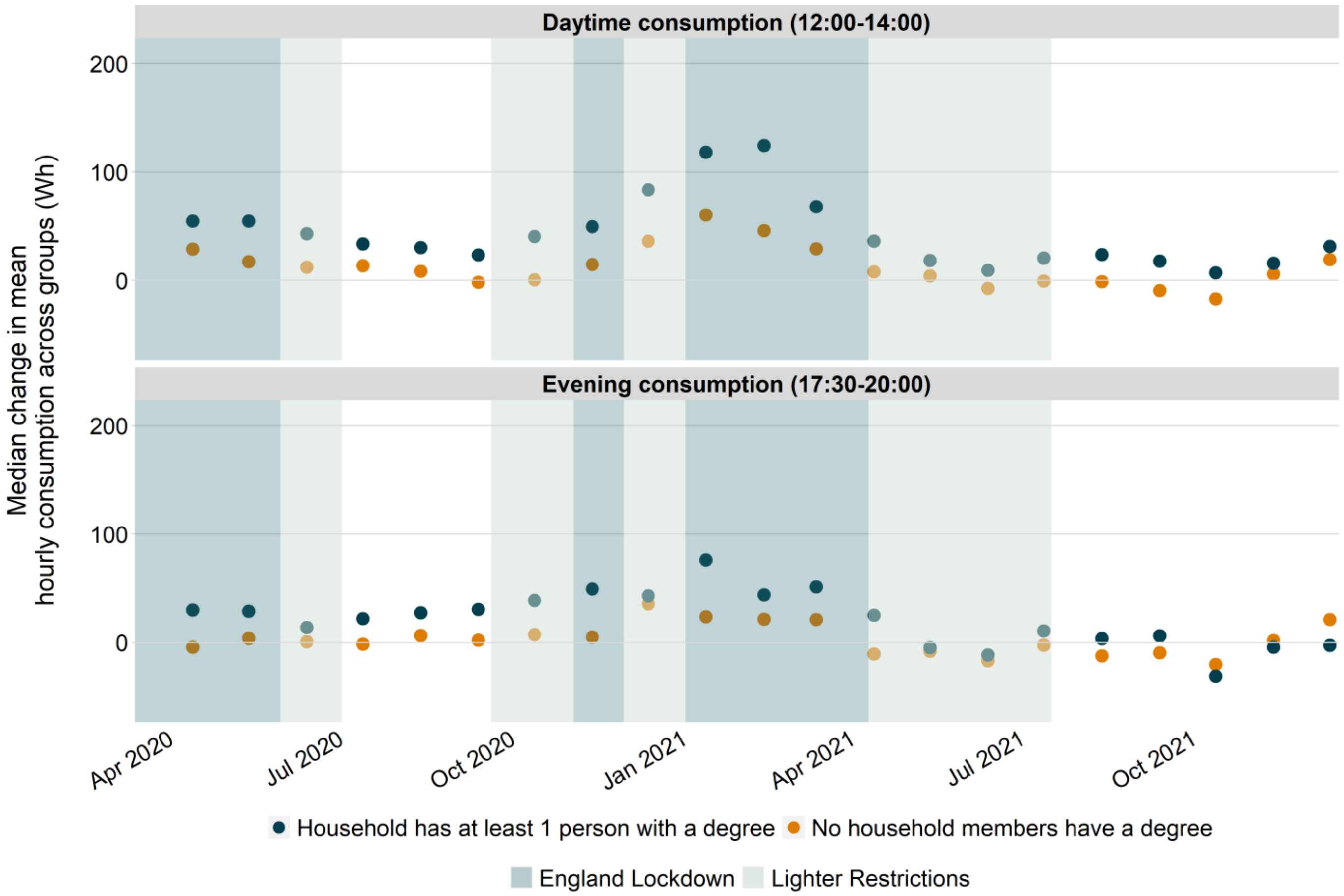
Households where at least one occupant has a degree

As part of considering households with different employment status and home-working ability, we consider households where at least one occupant has a degree. We consider this to be a proxy for households where occupants have office-based jobs, and are therefore typically more likely to be able to work from home.[26]

[25] This is discussed in more detail in Annex 4.

[26] As shown in Figure 46, households that work from home are more likely to have at least one occupant with a degree.

Figure 17 Change in daytime and evening consumption – split by households with and without at least one occupant who has a degree



Source: *Frontier Economics / UCL analysis using SERL data*

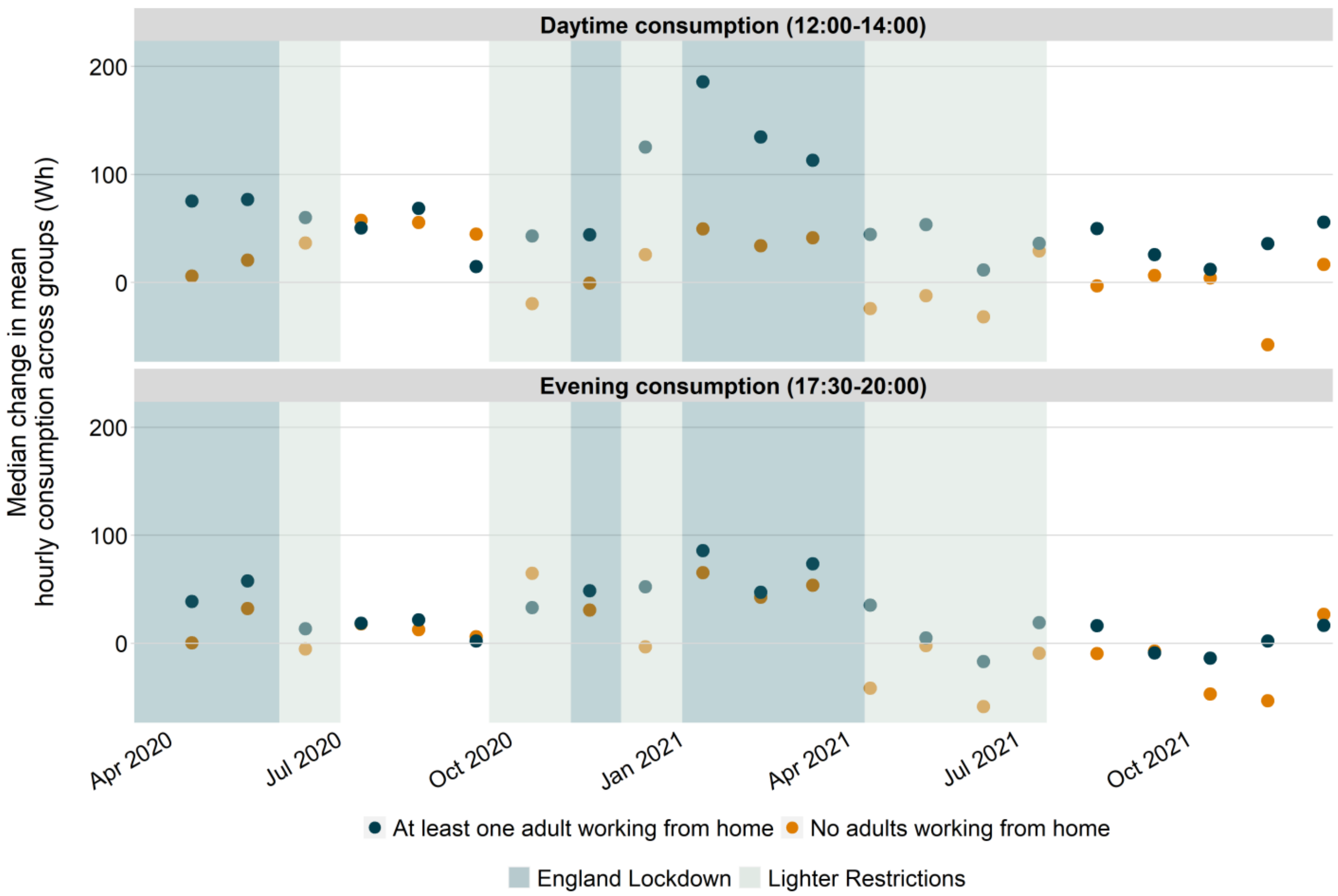
The change in daytime consumption tended to be positive across all time periods for households where at least one occupant has a degree. The largest changes in consumption tended to be during the lockdown periods; while periods outside of lockdown saw much smaller changes in consumption. However, the changes in consumption were most pronounced during the winter 2020/21 lockdown for households where at least one occupant has a degree.

The change in evening consumption tended to be much lower than the corresponding change in daytime consumption across all months for households where at least one occupant has a degree. During the lockdown periods, households where at least one occupant has a degree changed consumption more than other households, but by the end of the period there is little difference in changes in consumption. By the end of the period, the change in consumption was very similar across household types and close to zero.

COVID-19 survey on homeworking

Homeworking can also be measured by identifying households that responded to a survey stating that they would be working from home. The SERL COVID-19 survey[27] asked a question regarding whether participants were working from home often or occasionally. For households that have at least one occupant stating they would be homeworking either "always" or "sometimes", we consider these as homeworkers. However, only a subset of the total sample answered these questions, therefore this analysis is based on a reduced sample size[28] whilst the participants who answered the questions may in some way bias the survey outcome.[29]

Figure 18 Change in daytime and evening consumption – split by households with and without at least one member who worked from home



Source: Frontier Economics / UCL analysis using SERL data

[27] This survey was sent out and completed by participants in May/June 2020.

[28] Approximately 62.6% of our sample answered the COVID survey.

[29] For example, there could be a reason or common factor why these households responded to the COVID survey and why other households didn't. Related to this, there may be types of households that answered the survey question that have an inherent characteristic that is correlated with households where at least one occupant works from home. Therefore, the sub-sample of households that answered the SERL COVID survey may not be representative of the overall sample of households.

The change in daytime consumption diverged between the types of households for both lockdown periods in spring 2020 and winter 2020/21, as well as in November and December 2021. For households where no occupants work from home there were few positive changes in consumption across the time period. For households that work from home, there were large positive changes in consumption across most of 2020, winter 2020/21, and certain months across 2021 (see Section 4.1 and Annex 3 – changes in consumption for charts of the change in consumption for different periods of the day and estimated COVID-19 cases in England).

The change in evening consumption typically shows smaller positive changes in consumption and larger negative changes in consumption for both types of households across the time period. Both types of households show similar changes in evening consumption across the time period, although during lockdowns households where at least one occupant works from home tend to have slightly larger changes in evening consumption than other households. It is also worth noting that there are large decreases in evening consumption in some months where no adults are working from home. There are a number of factors that may be driving this, for example it may be that these workers face changing working patterns, or that they are socialising more as lockdown restrictions ease.

Implications for our hypotheses

The analysis above is consistent with our hypotheses that households working from home increased their daytime and evening consumption, typically during the lockdown periods. The largest change in consumption was, as expected, during the daytime in the lockdown period.

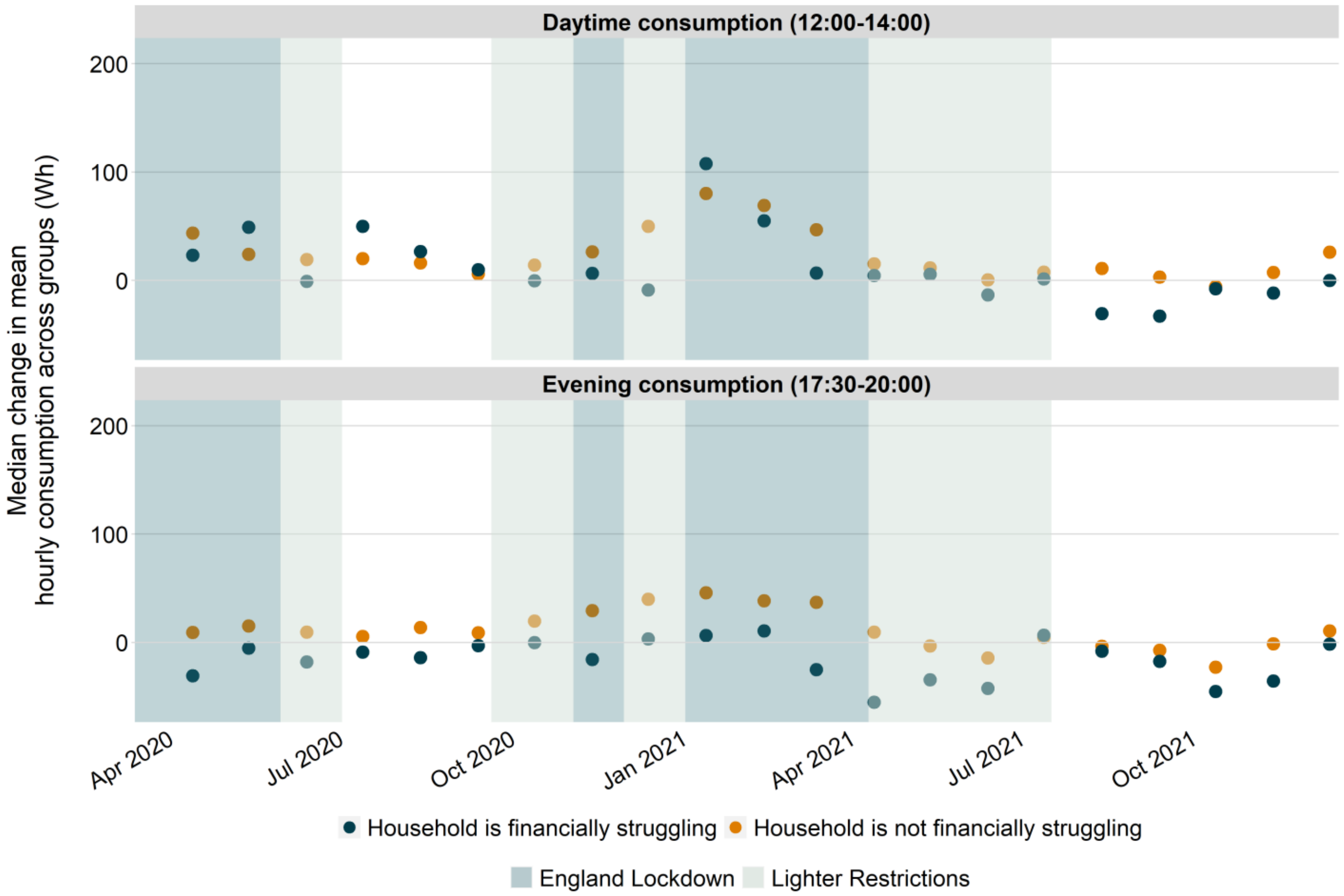
The analysis does suggest some persistence in behaviour for home-workers, although the effect is somewhat weakened compared to during the peak of the lockdown period. Ideally, the next step would be to identify the types of home-working households that continue to work from home in the future. For example, we want to understand whether it is younger or older households, or whether it is those on higher or lower incomes. We explore this analysis, and its limitations, in Section 4.

Income

How well are survey respondents managing financially

We now consider how differences in household income might affect changes in electricity consumption. We want to understand whether customers in vulnerable situations, such as customers on low incomes, have different responses to the pandemic compared to other customers. Whilst we do not have specific income data for households in the SERL data, there is a survey question regarding whether customers are 'managing financially'. We define customers who rate themselves as not managing well financially as 'financially struggling'.

Figure 19 Change in daytime and evening consumption – split by households which are and aren't 'managing financially'



Source: Frontier Economics / UCL analysis using SERL data

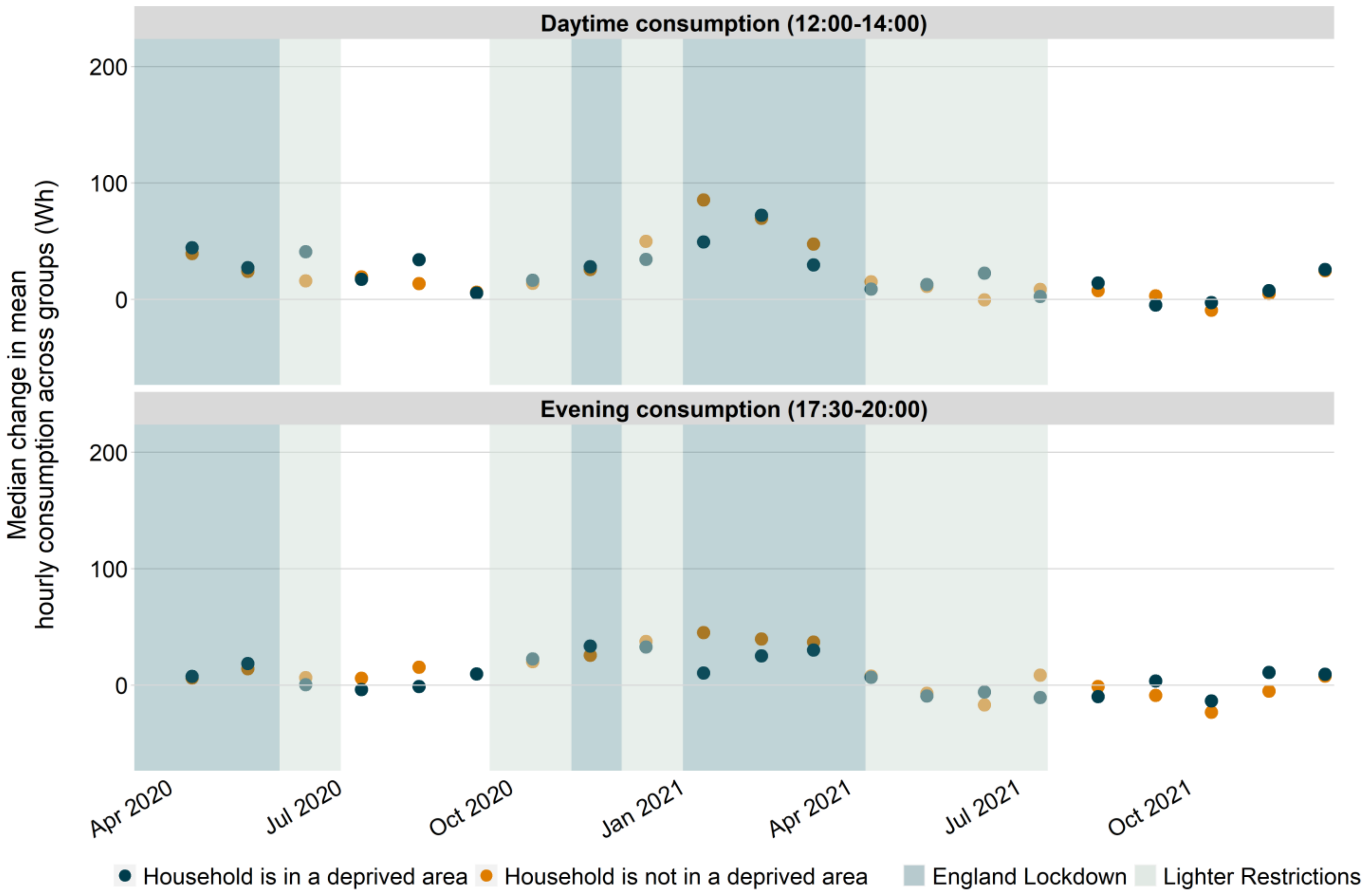
The change in daytime consumption showed a mixed picture between the two household groups. There was not a clear pattern of one group changing demand more consistently than the other group. During 2020, in some months households that are financially struggling showed greater change in consumption, whilst in other months households that are not financially struggling showed greater change. However, in 2021, households that are not financially struggling tended to show slightly larger changes in consumption than households that are financially struggling.

The change in evening consumption showed a more distinct picture between the two household groups. Households that are not financially struggling tended to have a small but positive change in consumption; whilst households that are financially struggling tended to have a small but negative change in consumption. This pattern for financially struggling households occurred for some months of the spring 2020 lockdown, the lockdown of winter 2020/21, and the spring of 2021. There were also some negative changes in consumption during the autumn of 2021.

Index of multiple deprivation

Another way of considering differences in financial wellbeing is identifying the prosperity or deprivation of the household's local area. We might expect households located in areas of relative deprivation to be financially struggling more than households located in areas of relative prosperity. We consider households located in areas[30] that are in the bottom quintile of the Index of Multiple Deprivation[31] as financially less well off. There are drawbacks with this approach, not least that that there is typically large variation is financial status within a local area and that the Index of Multiple Deprivation considers other aspects aside from financial wellbeing.[32]

Figure 20 Change in daytime and evening consumption – split by households which are and aren't located in a deprived area



Source: Frontier Economics / UCL analysis using SERL data

[30] Lower Super Output Areas (LSOAs)

[31] English indices of deprivation, 2019, MHCLG, https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019

[32] The Index of Multiple Deprivation proxies for more than just financial wellbeing, and therefore might be a more well-rounded representation of vulnerabilities. However, it is a wide-ranging description, some of which might not be categorised as vulnerability, such as the 'living environment'.

Changes in daytime and evening consumption were broadly similar for both types of household across all time periods. There were very few months that indicate any differences in consumption and in those months where there were some differences, there was no distinct pattern.

Conclusions

Our analysis of changes in consumption by customer group suggests the following.

- **The largest changes in electricity consumption typically occurred during periods of lockdowns and during the daytime rather than evening.** This is consistent with our analysis of overall changes in consumption (see section 3.2.4). This occurred for retired households, households where at least one occupant has a degree, households with at least one occupant working from home. Households with children had large increases in daytime and evening consumption. The large increase in evening consumption for this group was not one of our hypotheses.
- **Households with at least one child and those more likely to work from home had positive changes in consumption outside of periods of lockdown.** This suggests that behaviour changes during lockdowns which affected consumption for these groups persisted to some extent when lockdowns were lifted.
- **Retired households had no or very small changes in consumption in non-lockdown months**. This suggests that behaviour changes during lockdowns which affected consumption did not persist when lockdowns were lifted for this group.
- **Customer groups linked to lower incomes exhibited smaller changes in consumption.** Changes in consumption were similar for households regardless of whether they were located in a deprived area, and other changes in consumption are negligible beyond increased daytime consumption during periods of lockdown. We observe the same patterns for households that are financially struggling.

3.3.2 Changes in consumption by clustering

Introduction to clustering

The previous sections consider how household electricity consumption changed, first at an aggregate level, and then at a disaggregated level by dividing customers into groups using observable characteristics. Each of these approaches are useful and reveal information about household consumption changes. However, the aggregate analysis may hide the individual distribution, due to different types of behaviour. Not everyone in the sample is affected in the same way. When we divide customers into groups using observable characteristics, there may be more similarities in the way households in these groups respond, but there are still likely to be varied behavioural responses within these broad groups. We now examine the change in consumption for groups of households that exhibit similar behaviour.

Ideally we would do this by looking directly at the raw data. However that is not possible due to SERL data restrictions.

Instead, we use clustering analysis to identify groups of households with similar consumption behaviour, take the average changes in consumption within those groups, and consider those changes in consumption as representative of households in that group. For more detail on how clustering works, see Annex 4 – Clustering analysis. In our analysis, we are interested in grouping customers together that show similar behaviours in their daytime and peak changes in consumption, as well as customers that show similar changes over time. Therefore, we consider the changes in consumption for 12:00-14:00 and 17:30-20:00 as the relevant periods of consumption. Given that we are interested in changes in consumption, we adjust average consumption levels for each household to the Typical Domestic Consumption Value (TDCV)[33]. Other data cleaning and adjustment steps are set out in Annex 4 – Clustering analysis.

It is worth noting that we map the clusters to household characteristics to understand how many people in a given cluster have a specific characteristic. However, given that some clusters have small samples, dividing them into smaller sub-samples would have been potentially disclosive[34]. Therefore, to err on the side of caution, we have chosen not to map the clusters onto all characteristics listed above. Instead, we map onto two characteristics:

- households with a degree[35], and
- households with retirees.

We first set out below our hypotheses for what we expect to see as part of the clustering analysis. The hypotheses in this section do not necessarily link to households with a specific characteristic; rather we consider what might happen to different groups of households over time, given that we know different groups of households have different behaviours. Building on the aggregate hypotheses set out in the previous section, we consider the following changes in consumption.

- Some subsets of customers change their **daytime consumption** in response to the pandemic, and over the period there are likely to be different responses:
  - some customers may maintain the change in daytime consumption over the whole period, signalling that the pandemic has permanently shifted their behaviour;
  - some customers may maintain the change in daytime consumption over the lockdown periods, but revert back towards expected consumption in the non-lockdown periods; and
  - some customers may not change daytime consumption at all over the whole period.
- Some subset of customers change their **evening consumption** in response to the pandemic, and there may be some relationship between customers changing evening consumption and daytime consumption:

[33] The TDCV (as set by Ofgem) for electricity is 2,900 kWh for 2021. This is set out in the 'Typical Domestic Consumption 2021 Decision Letter': https://www.ofgem.gov.uk/sites/default/files/docs/2021/05/tdcv_decision_letter_2021_0.pdf

[34] Disclosure refers to someone being able to use statistical information, such as data, to infer confidential information, and potentially identify people or entities, from results, such as charts. See: https://ukdataservice.ac.uk/app/uploads/thf_datareport_aw_web.pdf

[35] Whilst we are primarily interested in working from home, we cannot use the working from home variable from the COVID survey due to the risk of disclosure. We therefore use households where at least one member has a degree as a proxy. As discussed in Figure 46 there is a positive correlation between households where at least one occupant has a degree, and households with at least one person working from home "all the time" or "sometimes".

  - some customers change both evening and daytime consumption at the same time (this might be a permanent or temporary shift in daytime and evening consumption);

  - some customers change daytime consumption but not evening consumption, suggesting only daytime behaviour has shifted;

  - some customers do not change either daytime or evening consumption.

- We also consider the subset of customers that show the largest changes in consumption during the lockdown period. Our hypotheses suggest that household composition may have changed, particularly during the lockdown periods (for example, where people in single-occupancy households moved in with partners or families). Therefore, we would consider that large changes in consumption do not correlate with persistence. Conversely, we should expect to see drops in consumption when the number of household members decreases, especially during periods of restrictions. Whilst the analysis of changes in consumption by customer characteristics (in Section 3.3.1) suggests that households with some characteristics saw decreases in consumption, this typically didn't occur during periods of lockdown.

In the following section, we have considered clustering over changes in daytime consumption and changes in evening consumption across six months from the sample. We also considered other types of clustering analysis and these are described in more detail in Annex 4 – Clustering analysis.

Clustering results

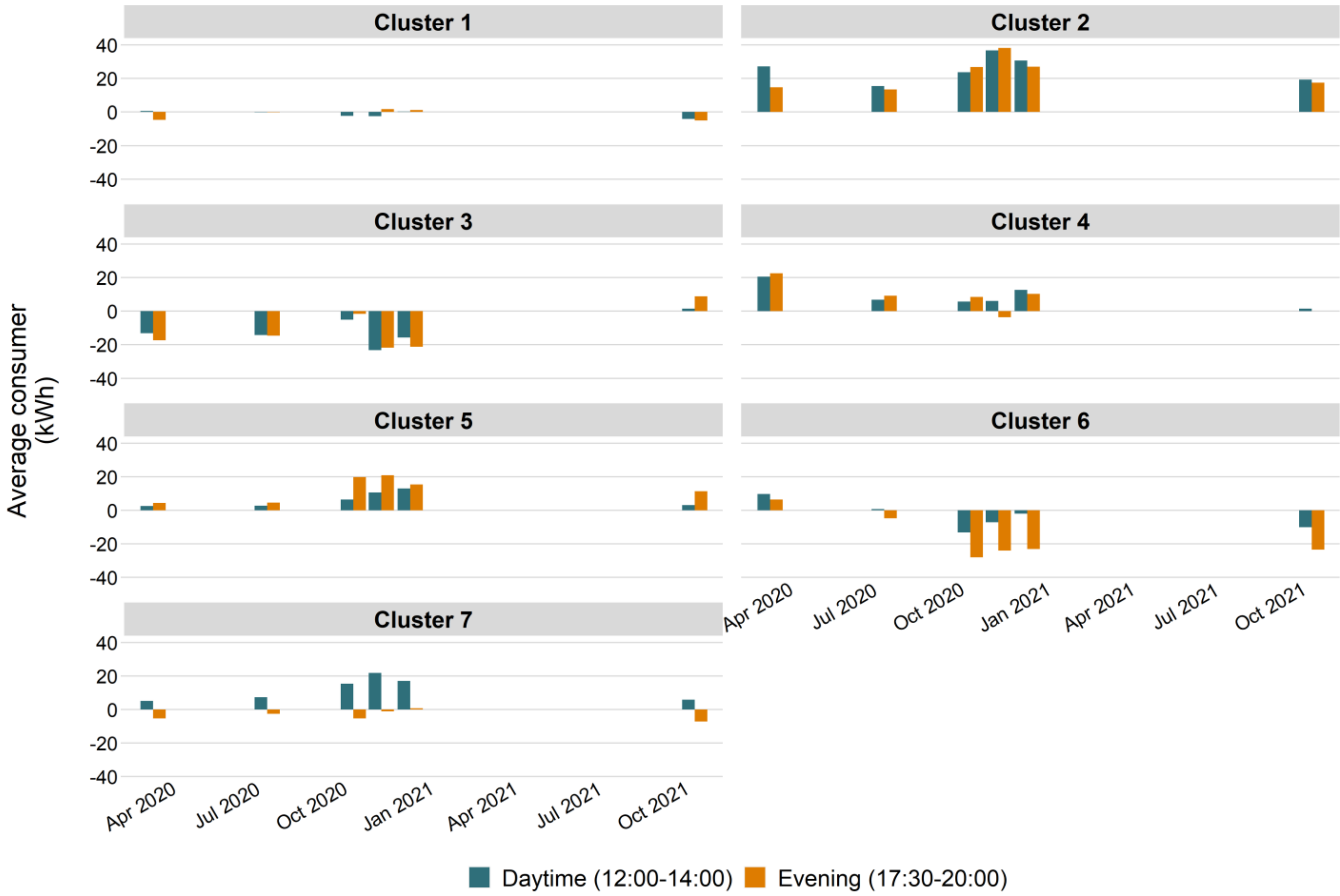
The idea of this analysis is to group households into consumption groups based on patterns of change in daytime and evening consumption, and how those vary over time. We choose six months to cluster over. These are: April, August, November and December 2020; and January and November 2021. We chose these six months to reflect various periods of lockdown, other restrictions and non-lockdown. This mix of periods might help reveal groups of customers with similar behavioural patterns over time.

Figure 21 below shows the change in consumption by cluster when the number of clusters are set to seven[36]. By imposing seven clusters, different patterns start to emerge and we can identify which groups are most prevalent[37]. We have also run the clustering analysis with the number of clusters set to two. When we set the number of clusters to two, there is a group of customers that tend to change consumption, and there is a group of customers that do not change consumption. The results of the two cluster analysis are discussed in more detail in Annex 4 – Clustering analysis.

[36] Although the Silhouette Method suggests that 2 is the optimal number of clusters for the analysis undertaken for multiple months, we present here the results when we set the number of clusters to 7. As can be seen in Figure 57, there is a spike when 7 clusters are evaluated, and 7 clusters provides more granular and different clusters of household changes in consumption.

[37] For the seven cluster analysis, we do not display data on the proportion of households within each cluster that are retired or have at least one person with a degree. This is due to SERL's statistical disclosure control (SDC) policy, which requires that data points (such as a bar on a bar chart) are calculated using data from at least 10 households.

Figure 21 Changes in daytime and evening consumption over time by cluster – seven cluster analysis



Source: Frontier Economics / UCL analysis using SERL data

There are broadly three types of clusters.

1. **Clusters containing households that changed consumption and persisted (to some extent)**

- Cluster 2 (8% of the sample) are households with large, positive changes in daytime and evening consumption throughout the pandemic, including the periods outside of lockdown (such as August 2020 and November 2021).
- Cluster 5 (20% of the sample) are households with positive changes in daytime and, more predominately, evening consumption with the largest changes during winter 2020/21. The small changes in consumption in November 2021 suggest persistence is limited.
- Cluster 6 (9% of the sample) are households with positive changes in daytime and evening consumption in the initial period (April 2020) but negative changes in consumption, particularly in the evening, for subsequent periods.
- Cluster 7 (16% of the sample) are households with a positive change in daytime consumption, and small negative changes for evening consumption. The largest positive change in daytime consumption occurs

during winter 2020/21, whilst the negative changes are reasonably small across all time periods. The small changes in consumption in November 2021 suggest persistence is limited.

2. **Clusters containing households that changed consumption and did not persist**

- Cluster 3 (4% of the sample) are households with negative changes in daytime and evening consumption throughout the pandemic, apart from November 2021 where there is little change in daytime and evening consumption.

- Cluster 4 (13% of the sample) are households with positive changes in daytime and evening consumption in the initial period (particularly in April 2020) but tends to diminish over time, until November 2021 where there is no change in consumption.

Clusters containing households that did not change consumption

- Cluster 1 (31% of the sample) is made up of households that did not change daytime or evening consumption at all during any period of the pandemic. This is the largest sub-group in our analysis, and suggests a third of households may have seen electricity consumption unaffected by the pandemic.

This analysis shows that households had significant variation in consumption patterns throughout the pandemic. The nuances of these changes in consumption are hidden by more aggregated analysis; the clustering analysis helps to unpick these trends.

Implications for our hypotheses

Referring back to our hypotheses in Section 3.2, we find that the results of the clustering analysis support some of our hypotheses.

- There were a significant number of customers that changed daytime consumption in the initial period and subsequent lockdown periods. Some customers persisted with these consumption changes through to the end of the research period, such as cluster 2, whilst other customers reverted back to pre-pandemic consumption patterns by the same time. There were many customers that never changed daytime consumption at all, such as cluster 1, or temporarily shifted during the most severe lockdown periods, such as clusters 5 and 7.

- Changes to evening consumption followed a similar pattern to changes in daytime consumption for a majority of households, with households in clusters 6 and 7 being the exception. There were some changes that persisted (clusters 2 and 6), some changes that did not persist (cluster 4), and some households that did not change consumption at all (clusters 1 and 7). There were also broadly positive correlations between changes in daytime and evening consumption: households with positive changes in daytime consumption also tended to have positive changes in evening consumption. However, there were some households that showed positive changes in one period and negative changes in the other period (for example, cluster 7).

- It is a mixed picture regarding whether initial large changes in demand are correlated to subsequent persistence in consumption changes. There are some customers that show a positive change in consumption for both daytime and evening consumption in April 2020, and maintain this throughout the period. However, there are slightly more customers that show a similar positive change in consumption for both periods in April 2020, but do not persist in that behaviour.

3.4 Impact of the pandemic on customers and the network

So far, the analysis has described electricity consumption for specific customer groups. While that can reveal information about the differences in customer behaviour, we are particularly interested in how changes in electricity consumption might impact electricity DNOs. We consider two ways in which a change in electricity consumption might impact DNOs.

- First is the impact on the network, and particularly if there are changes to peak demand that might impact network capacity requirements. If we find the pandemic has increased peak demand (or decreased peak demand) then electricity networks might reprioritise planned network build, or reconsider whether further expansion is required.

- Second is the impact on customer bills, and particularly for vulnerable customers in fuel poverty. All DNOs have increased responsibility for customers in fuel poverty throughout ED-2[38], so it is important to examine whether they have been impacted by the pandemic, and estimate the potential size of the impact.

We consider each of these in turn.

3.4.1 Network impact

As highlighted above, DNOs plan their network capacity around demand. In light of large shifts in consumption behaviour as a result of the pandemic, we want to understand how peak demand has changed. To do this, we consider the impact from average peak consumption increases from customers. We also contextualise these increases – are they significant given available headroom on the networks?

To do this analysis, we consider the percentage increase in domestic peak demand consumption[39] during the pandemic, and then identify how many locations on the distribution network might be impacted. We used the network capacity map[40] to identify the demand headroom[41] as a percentage of total capacity at primary substations. This can be thought of as the percentage increase in demand required to reach capacity at that substation. We then compare this to the average increase in demand.

Below, we show the results on a histogram in Figure 22. The red bars indicate the number of substations: substations with less demand headroom are on the left, while substations with more demand headroom are on the right. The thin blue line indicates the average change in peak demand domestic consumption at the height of the pandemic (around 5% increase).

[38] Ofgem, Consumer Vulnerability Strategy 2025, https://www.ofgem.gov.uk/publications/consumer-vulnerability-strategy-2025

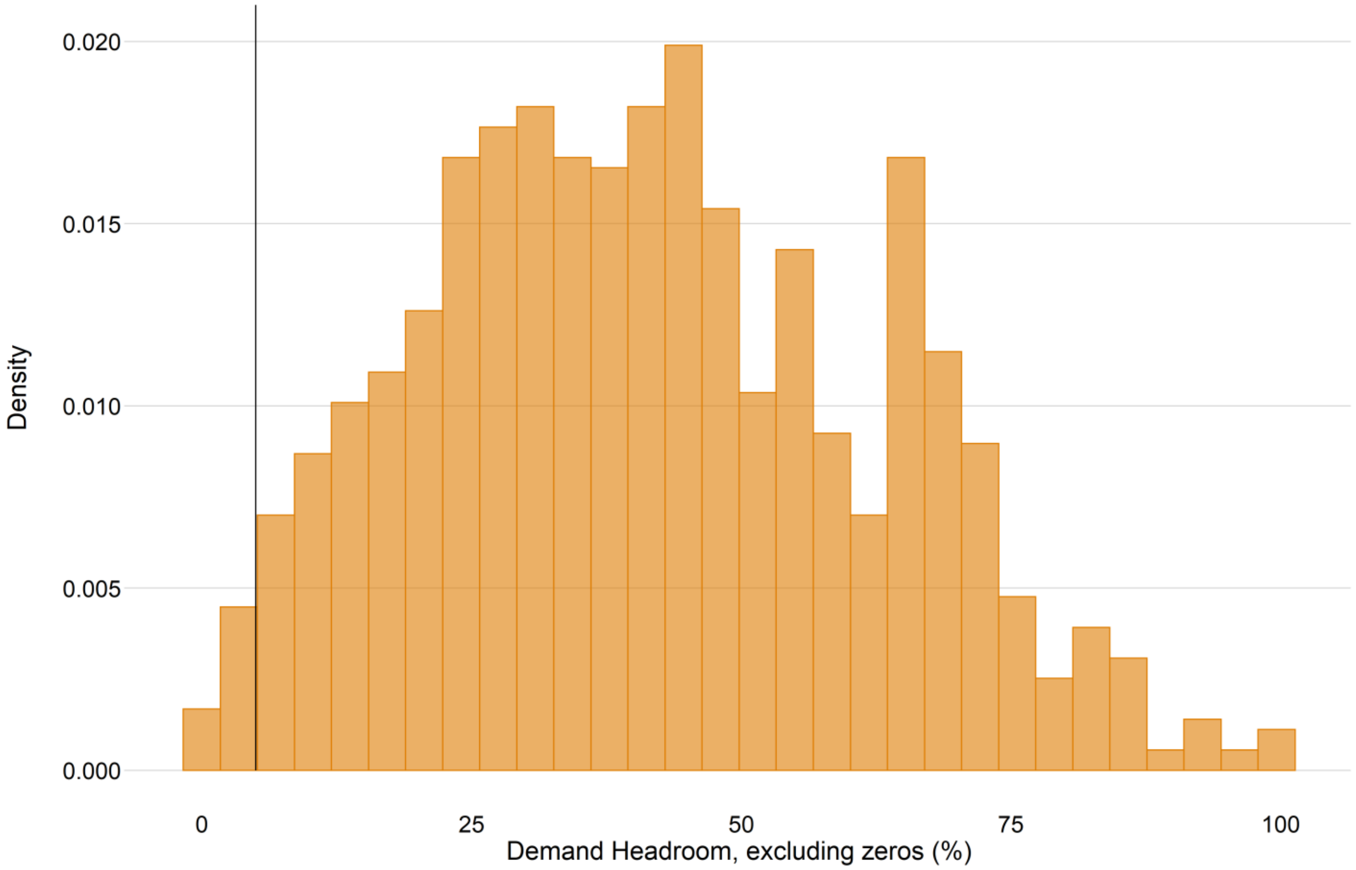
[39] Peak demand consumption is defined as average consumption during the evening peak (between 17:30 and 20:00) on weekdays in January.

[40] https://www.nationalgrid.co.uk/our-network/network-capacity-map/

[41] Where 'demand headroom' is defined as the difference between the ability of the electricity network to supply electrical demand, and the actual peak demand at that part of the network.

The substations to the left of the thin blue line are the ones with lowest demand headroom, and therefore might have been affected if the pandemic effect on peak demand persists into the future. The results below show that 2% of networks have a capacity of below 5%, where the pandemic may have taken them over their capacity. This suggests that the impact of the pandemic on the network is likely to be low in aggregate, but further investigation may be required on a small number of substations. We consider these in greater detail in the next section when considering the persistence of customer behaviour, and what that might mean for the network.

Figure 22 Only a small number of sub-stations are likely to be affected by an increase in peak demand like the increase seen during the pandemic



Source: Frontier Economics/UCL/NGED network capacity

Note: Black line indicates the average increase in winter peak demand during the pandemic; ~5%.

3.4.2 Vulnerable customer impact

The second impact that we consider as a result of the pandemic is the impact on customer bills. One result from the previous section is that there is a subset of customers who increased electricity consumption throughout the pandemic period. One of the impacts of increased consumption for these customers is (all else being equal) a larger customer bill. We can use these estimates of consumption changes to estimate the range of customer bill impacts.

The pandemic had a wide variety of impacts on households, and there were large differences in cost impact of the pandemic on the richest and poorest for goods and services other than energy, with the poorest tending to have more severe impacts.[42] NGED has a commitment to provide support for its most vulnerable customers, with Ofgem stating that “we want to continue to improve the service provided by distribution network companies to consumers in vulnerable situations including fuel poor”.[43] Therefore, it is important for NGED to consider whether the pandemic exacerbated this cost impact on poor customers, and for NGED to identify the specific type of vulnerable customer that was impacted by the pandemic.

While we do not have income data that we can match to household consumption data, the accompanying survey data does include self-reported information on whether a customer is “financially struggling” or not. We define financial wellbeing as “good” if they selected “living comfortably” or “doing alright” and as “poor” if “just about getting by” or “finding it quite difficult” or “finding it very difficult”. Figure 23 below shows how monthly consumption changed by financial wellbeing status.

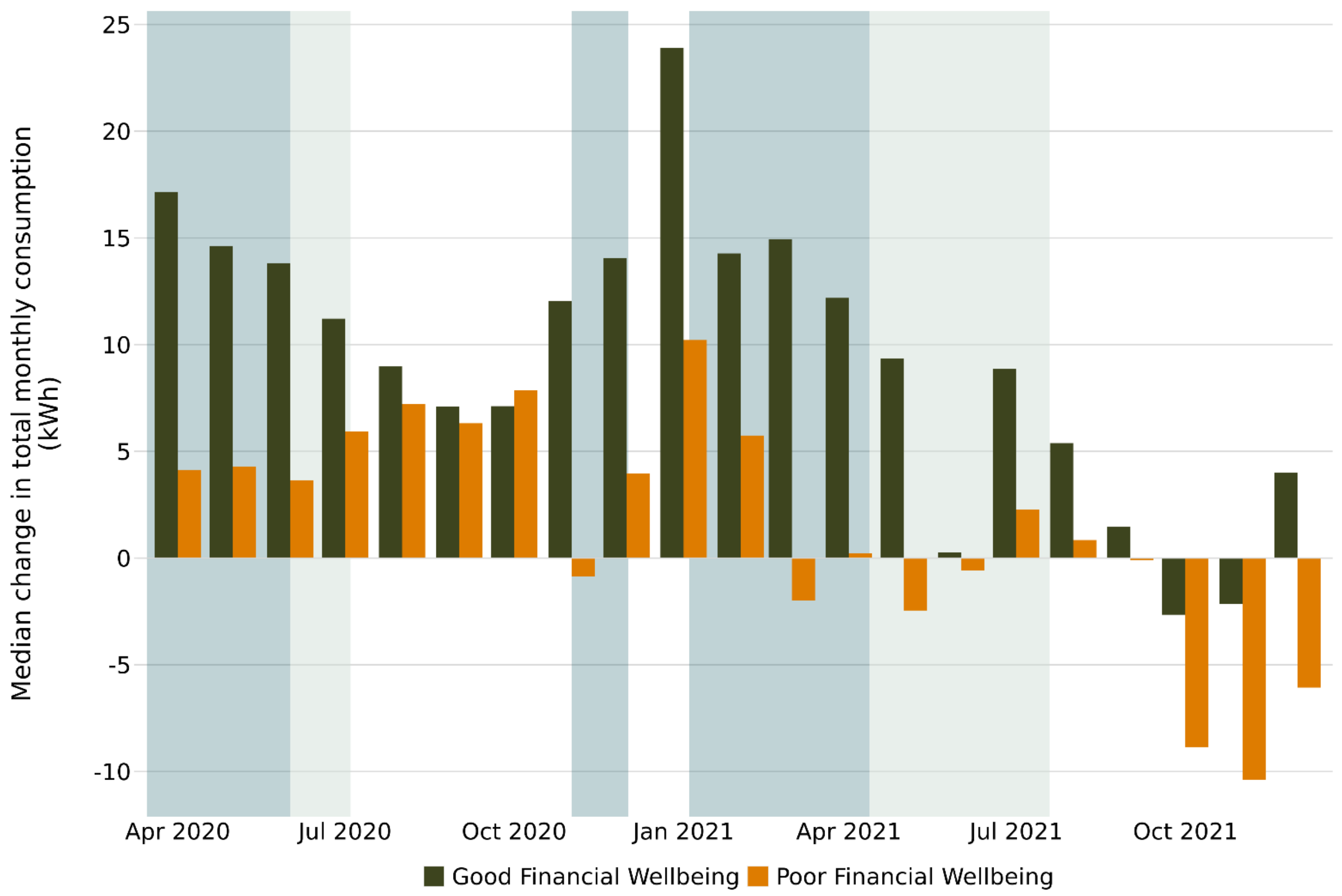
It should be noted that, outside of the lockdown periods, there was little difference in the change in consumption across households with different financial wellbeing[44].

[42] Cantó, O., Figari, F., Fiorio, C.V., Kuypers, S., Marchal, S., Romaguera-de-la-Cruz, M., Tasseva, I.V. and Verbist, G. (2022), Welfare Resilience at the Onset of the COVID-19 Pandemic in a Selection of European Countries: Impact on Public Finance and Household Incomes. Review of Income and Wealth, 68: 293-322. https://doi.org/10.1111/roiw.12530

[43] Ofgem, Consumer Vulnerability Strategy 2025, https://www.ofgem.gov.uk/publications/consumer-vulnerability-strategy-2025

[44] At first glance, Figure 23 may look inconsistent with Figure 19, which suggests that financially struggling households have no change or a negative change in consumption for many months. However, this may be due to Figure 23 including all periods across the day, whereas Figure 19 only looks at snapshots (daytime and evening) from the day.

Figure 23 Households with poor financial wellbeing were less likely to increase their electricity consumption over the pandemic period



Source: Frontier Economics / UCL analysis using SERL data

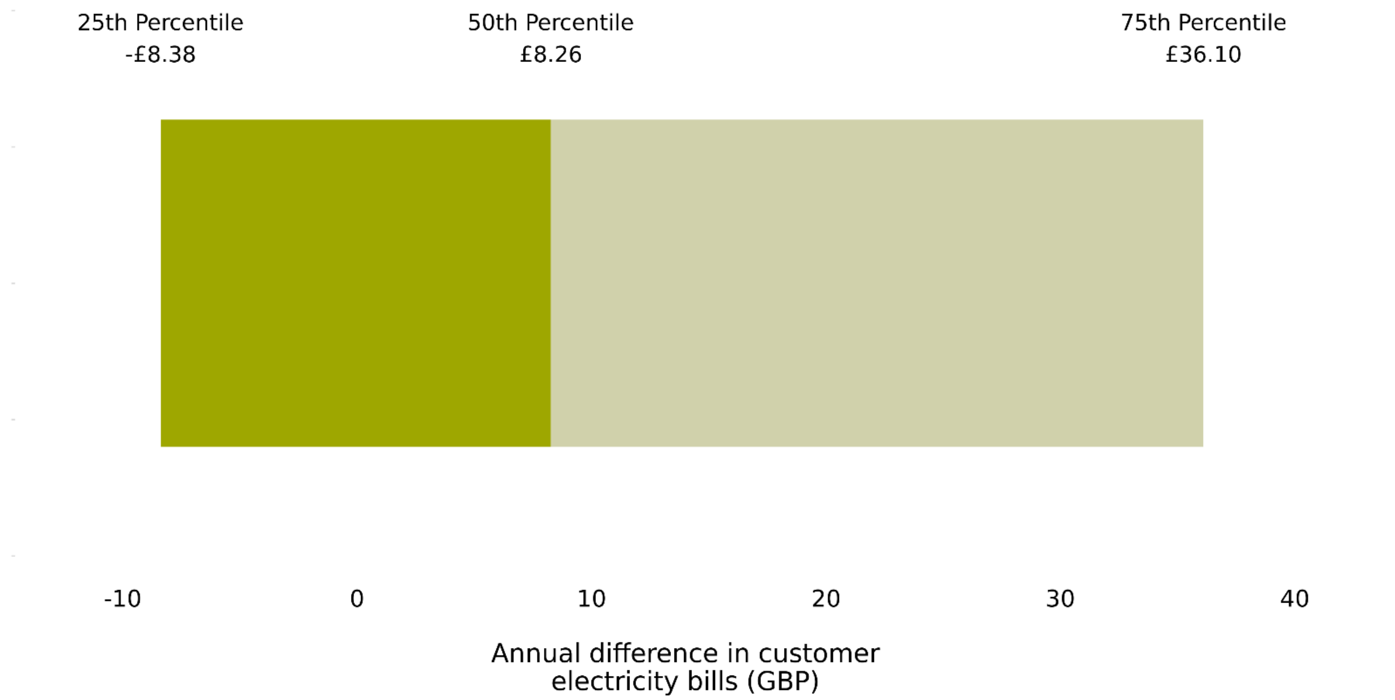
Note: Changes in total monthly consumption by self-reported financial wellbeing. We class financial wellbeing as "good" if they selected "living comfortably" or "doing alright" and as "poor" if "just about getting by" or "finding it quite difficult" or "finding it very difficult".

Before we consider the impact of the pandemic on vulnerable customer bills, we want to consider the estimated bill impact. We calculate this cost difference on the assumption that all customers were on tariffs with unit rates in line with Ofgem's default tariff cap. The cost is summed over a year, from April 2020 to March 2021 inclusive. This calculation corresponds to the changes in consumption outlined earlier in this section, converted into customer bills.

The results show that the median is a small increase of about £8 for that yearly period. Our analysis does only cover changes in electricity consumption, and so if gas consumption also rose then this could have led to a higher effect (particularly as some lockdowns coincided with winter, when the need for heating is highest). However the resulting increase would still likely be small, particularly when compared to the rises in bills caused by the increase in wholesale market prices over the course of 2022.

However the range of outcomes is broad: the top 25% of households (in terms of overall consumption increases) saw their bill increase by just over £36, while some households saved a small amount of money, reflecting their decreased consumption. This reflects the results of the changes in consumption by customer groups: different customer groups showed changed their behaviour in different ways as a result of the pandemic.

Figure 24 Total annual bills tended to increase during the pandemic (April 2020 – March 2021)

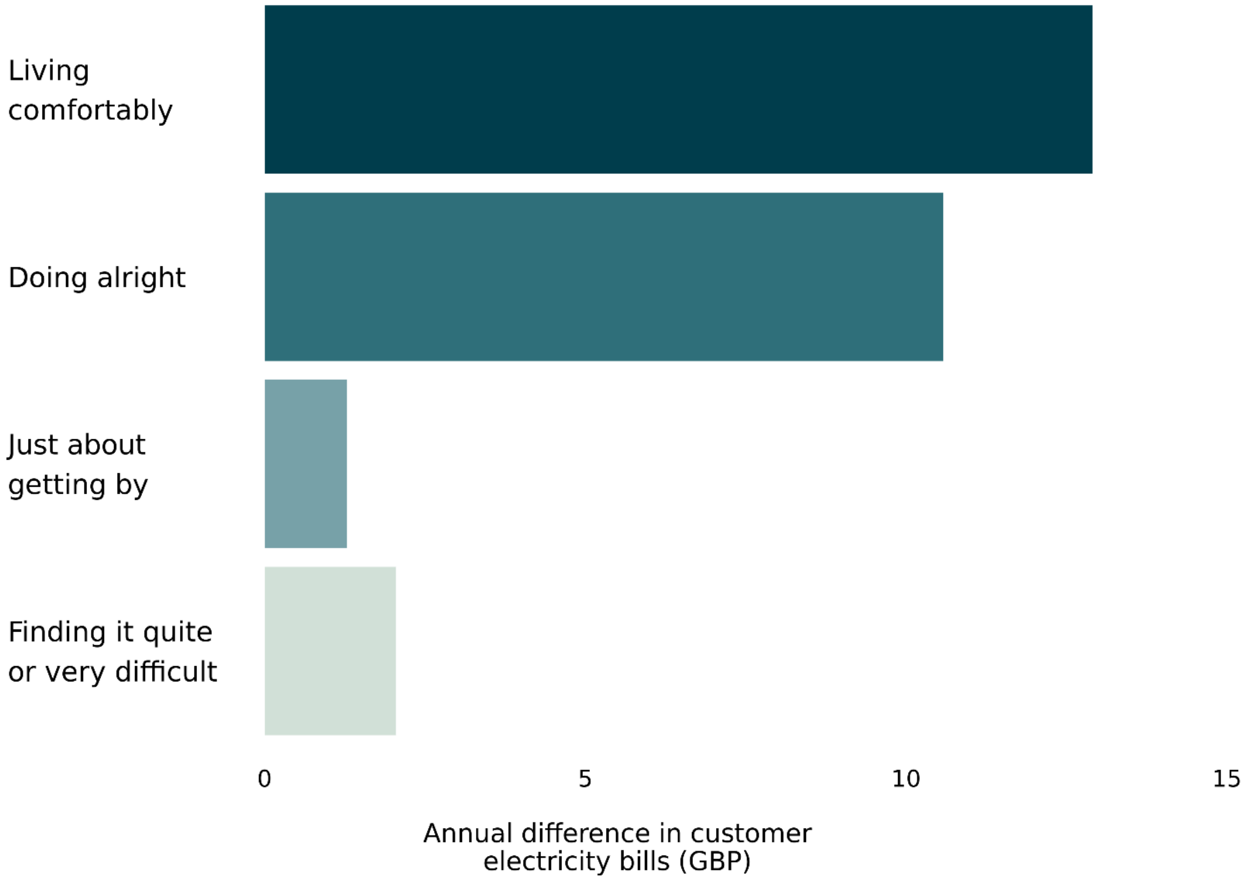


Source: Frontier Economics / UCL analysis using SERL data

Note: Percentiles are approximations to preserve data anonymity

We can also split the sample by how they answered the financial wellbeing question (not answered by all). Figure 25 shows a substantial difference between those who were doing better financially, increasing spending by around £11-£13 compared to those who were struggling financially, increasing by £1-£2 only.

Figure 25 Total annual bill change was larger for those with good financial wellbeing



Source: Frontier Economics / UCL analysis using SERL data
Note: Chart shows approximate median bill change, data are approximations to preserve data anonymity

This analysis suggests that, during the lockdown period, the pandemic had a much smaller impact on the electricity bills of those struggling financially than those managing well. This could be for three reasons:

- One reason might be because households struggling may have been in jobs that could not be done from home, hence no changes to their working location leads to little change in electricity consumption.
- A second reason could be because those with poor financial wellbeing are unable to work, and hence may have spent more time at home in the period before the pandemic. Therefore when the pandemic began, the consumption habits of those households would be unchanged.
- A third reason might be because those with poor financial wellbeing are less likely to have high energy usage. This could be because they do not own high energy-consuming appliances, or felt unable to use them due to cost.

3.5 Modelling conclusions

We can use this analysis to draw some conclusions on the impact of the pandemic on domestic electricity consumption in England.

The overall impact was substantial, particularly during the "lockdown" periods in spring 2020 and winter 2020/21, with an average increase in consumption of over 6% across all household in April 2020. Assuming this increase is the same across all households, and households consume around 9KWh per day[45], then that translates into a UK-wide increase of around 14TWh. However by the end of the period (December 2021) there was a negligible difference between observed and modelled consumption, which suggests (at a high level) few lingering impacts of the pandemic at an aggregate level.

However, we also test hypotheses relating to different impacts at a more granular level.

- Our hypotheses suggest there might be increased daytime consumption among particular sub-groups of customers. Our modelling found, at an aggregate level, the daytime consumption increases were positive and much larger than increases during the evening peak. During winter 2020/21, these changes reached over 15% in the midday to early afternoon period and remained positive (but reduced) through to the end of the period.

- Given the biggest change associated with staying at home was on daytime demand, we did not consider any hypotheses specific to evening peak demand. There were some increases in evening peak demand during the height of the "lockdown" periods, but these increases diminished over the time period.

These patterns were evident among our population sub-groups, with the largest increases occurring in daytime consumption and smaller increases occurring in the evening peak period.

Analysis of changes in consumption by customer group shows the following.

- Several types of customer saw the largest increases during the daytime and during periods of lockdowns. This is the case for working (non-retired) households, households where at least one occupant has a degree and households with at least one person working from home. Households with at least one child saw large increases in both daytime and evening consumption.

- Of these households, some exhibited persistence of lockdown behaviour (and thus changes in consumption). This included households with at least one child, households where at least one occupant has a degree and households with at least one person working from home.

- Customer groups linked to lower incomes exhibit smaller changes in consumption, and were less likely to display clear patterns of changes in consumption. The clearest pattern occurred for financially struggling households, which had a small and negative change in evening consumption for months in or close to periods of lockdowns.

Clustering analysis of changes in consumption provides these additional insights.

- There are broadly three groups of households: a group that does not change consumption, a group that changes consumption persistently and a group that temporarily changes consumption. However, some of the analysis suggests that each of these groups contains sub-groups of more distinct patterns of changes in consumption. This includes households which have a negative change in consumption, households which may only have a change in consumption for one period of the day, and households

[45] https://www.ofgem.gov.uk/publications/decision-typical-domestic-consumption-values-2020

that have a material increase in consumption during periods of lockdown, which persist only for some households, not all.

The impact on customers and the network reflects the above trends. The impact of the pandemic resulted in increased domestic consumption, but this was generally concentrated among daytime periods and for a sub-group of customers.

- The impact on customers is therefore varied: some sub-groups of customers would have seen large impacts on their bills, while other sub-groups would likely see negligible impacts (or even slight declines in some cases). For customers with low financial wellbeing, their bills increased less than the average customer; suggesting the impact of the pandemic did not have a disproportionate impact on vulnerable customers.

- The impact on the network may be fairly limited. The increase in peak demand for winter weekdays during the pandemic was around 5%. If that behaviour persisted, given the small number of areas with demand headroom less than 5%, the impact on NGED's network would not be widespread. It may be significant for these parts of the network; we will investigate these in Section 4.2.

In the next section, we consider how these modelling conclusions can be combined with the findings from Section 2 on persistence and habit formation to understand how consumption behaviour during the pandemic might persist into the future. We consider what impact that might have on NGED.

4 How might changes in electricity consumption persist, and what is the future impact for NGED?

In this section we consider how behaviours that led to changes in electricity consumption might persist beyond the end of the pandemic and across ED2. We will want to consider the extent of these behaviours and the subsequent effect on NGED. We assess this in two parts.

- First, to the extent changes in consumption occurred, how might they persist in sub-groups of customers? Can we identify the factors driving that change in behaviour?

- Then, once we have identified behavioural changes, what does that mean for NGED? Are there persisting impacts on vulnerable customers? Is there potential for specific areas of the network to be impacted?

4.1 What are the factors driving behavioural changes, and how might they persist?

Our analysis in Section 3.2 finds that the pandemic drove changes in household electricity consumption (both in overall terms, and to the profile of that consumption during 2020 and 2021. In this section we want to understand how behaviours formed during that period might persist into the future. To understand that, we need to draw on the persistence assessment work. In particular, we draw on Section 2.3 where we set out the trigger-action-reward process that tells us how likely it is an individual might form habits. We use that work, alongside the modelling results, to investigate how behaviours from the pandemic might persist, and how pervasive those behaviours are.

As we discuss above, our analysis has found that the pandemic drove changes in electricity consumption. As the severity of the pandemic grew and declined over the period, behaviour changed with it. This can be illustrated by Figure 26, which shows the relationship between COVID-19 cases (as estimated by the Office for National Statistics, using an infection survey[46]) and the change in electricity consumption (Figure 26 shows changes in evening consumption). We focus on the change in evening consumption as this is when peak consumption occurs, and thus when there's most stress on the network. We consider the relationship between COVID-19 cases and the change in daytime and daily consumption in Annex 3 – changes in consumption.

Figure 26 below shows how the relationship between cases and changes in electricity consumption varied according to different phases of the pandemic. We show these in the three different panels in the figure.

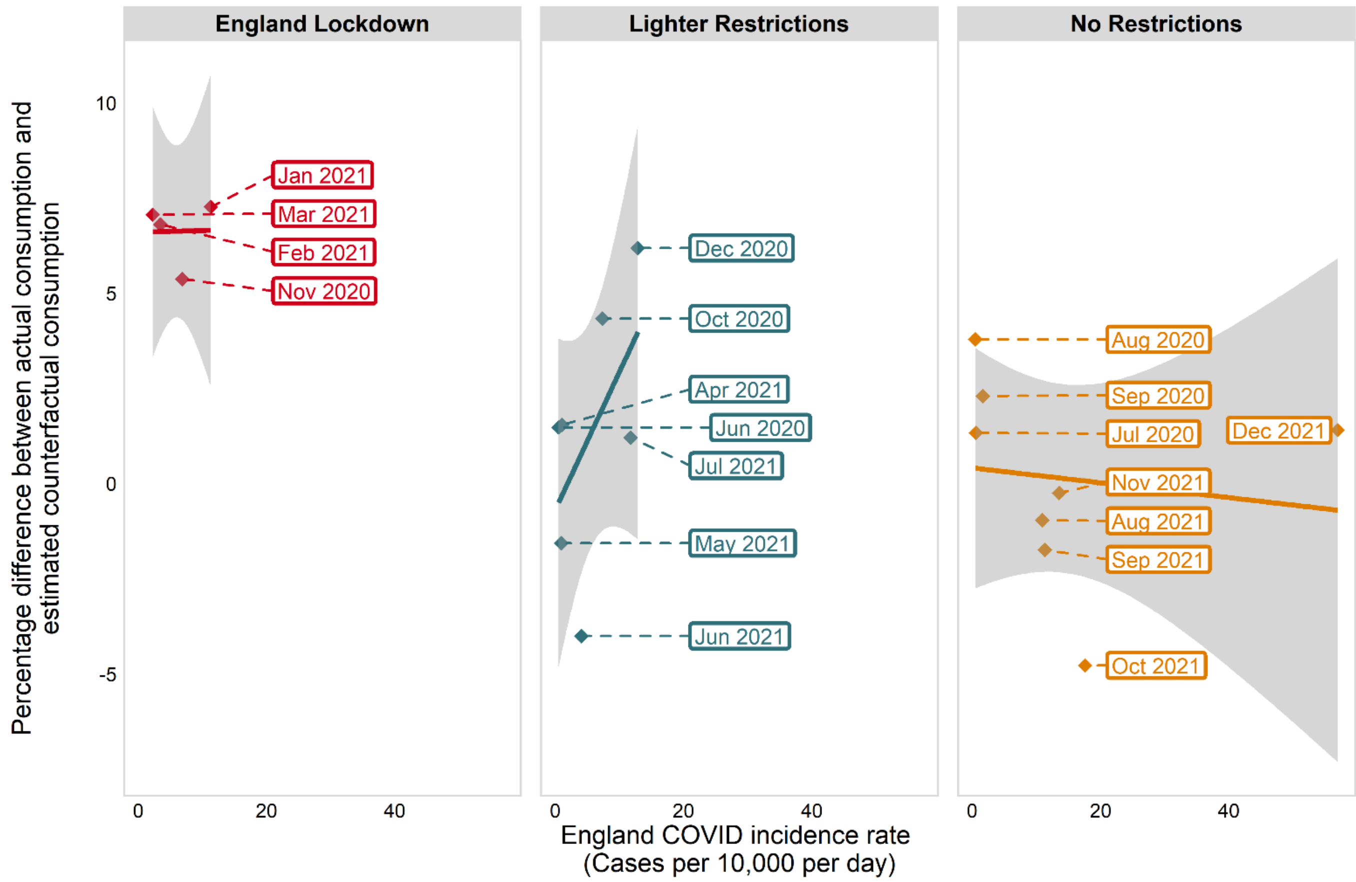
- **England lockdown**. During the lockdown periods, changes in consumption were larger, which shows how households changed behaviours in line with government restrictions.

[46] Office for National Statistics: Coronavirus (COVID-19) Infection Survey: England; https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/coronaviruscovid19infectionsurveydata

- **Lighter restrictions**. During the lighter restrictions, there was some correlation between case numbers and changes in electricity consumption. This could suggest households changed behaviour according to their perceived level of risk.

- **No restrictions**. Most relevantly for assessing the persistence of behaviour, during periods of no restrictions there was little correlation between cases and changes in electricity consumption. Therefore it might be the case that, in times of lighter restrictions when COVID-19 is on people's minds, they react to case numbers. But that relationship changes during times when restrictions were completely eased. Indeed, the relationship tended to reverse over time. In summer 2020, changes in consumption were larger than in autumn 2021 despite cases in the later period being higher (notwithstanding higher changes in December 2021). This suggests that, on average, households did not persist with behaviour changes during autumn 2021 that would have increased their consumption.

Figure 26 Relationship between estimated cases in England and changes in evening consumption



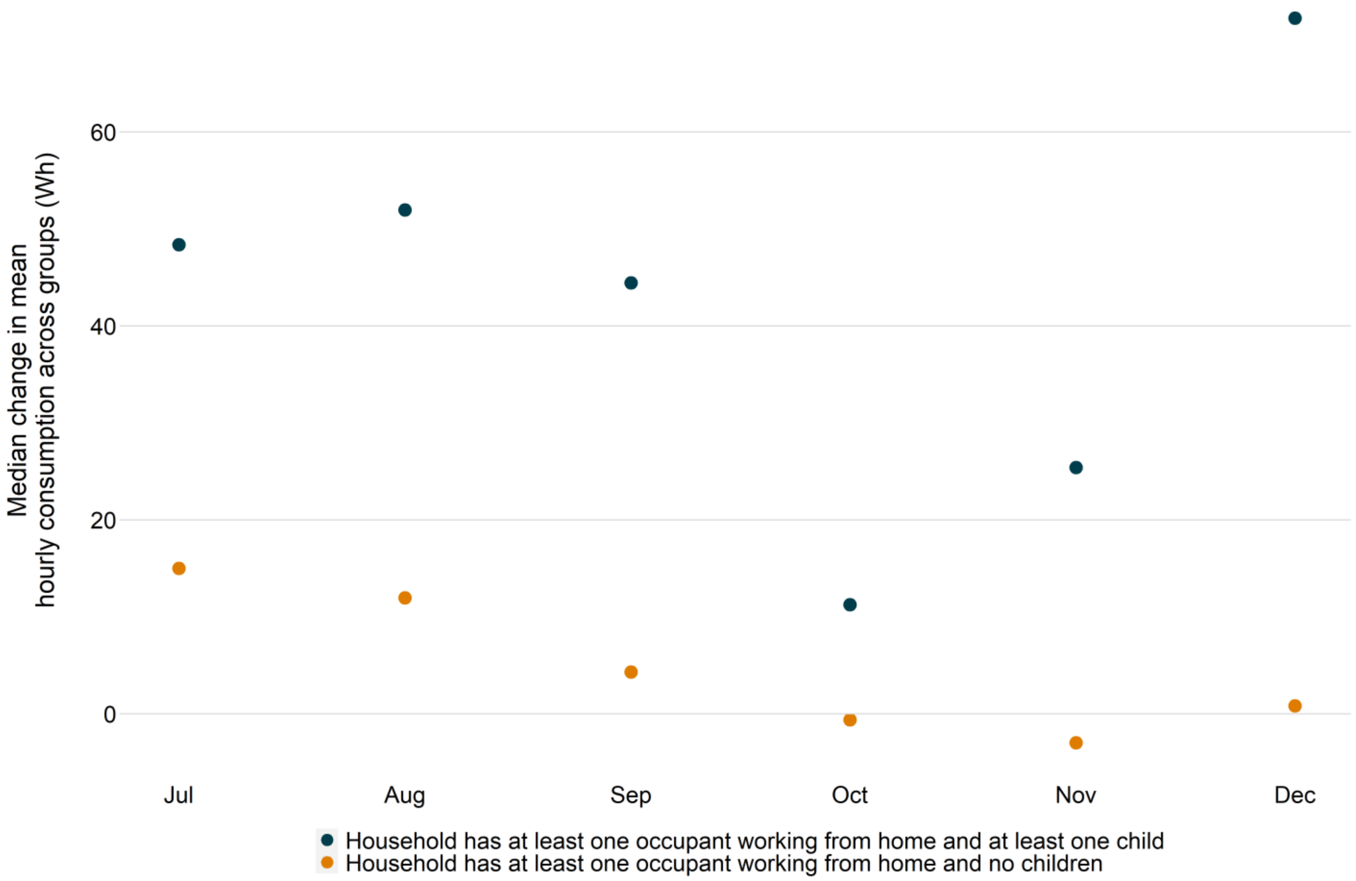
Source: Frontier Economics / UCL analysis using SERL data

Note: The grey shading shows the confidence intervals around the line of best fit. The wide shading indicates that these trends are not statistically significant.

To explore the underlying behavioural changes, we go back to the habit formation framework explored in Section 2.3. This framework suggests that the presence of a recurring reward can result in temporary behaviour changes turning into permanent changes. We identified a number of potential customer types that would receive rewards from remaining at home. Taking account the learnings from the modelling in Section 3.3 (considering the sub-groups of customers that tended to change consumption the most), we can reasonably conclude that households working from home and households with children are the two types of household that show most persistent behaviour.

We can test the extent to which households with both these characteristics tend to stay at home. Figure 27 shows that these households maintained large changes to their electricity consumption throughout the second half of 2021. The only outlier to this behaviour was in October 2021, which could be explained by these households taking advantage of eased travel restrictions and going on holiday[47].

Figure 27 Households that WFH with children changed their electricity consumption in the second half of 2021 more than other households



Source: Frontier Economics / UCL analysis using SERL data

So while we have not observed increases in domestic electricity consumption in aggregate as a result of the pandemic, there is some evidence of specific household groups having a sustained increase in consumption. Those with persistent changes in electricity consumption are likely to make a small proportion of the

[47] There were significant easing of travel restrictions, including merging of green and amber lists and changing testing requirements, https://www.caa.co.uk/news/2021-quarter-four-flight-data/

population: only 8% of our sample are households with children that indicated they work from home[48]. This is aligned with the results of our behavioural analysis in Section 2.4, which identified home workers as the most likely to persist in their behaviour, and identified a specific subset of those customers who are likely to receive some rewards for doing so.

However, as the effects of the pandemic appear to have declined over time, there are likely to be other factors that could outweigh any lingering persistence in behaviour. The rise in gas prices throughout the second half of 2021 and 2022 has led to large increases in domestic energy costs. It is unclear to what extent this change in costs may impact households in the future, and how this will interact with behaviours developed during the pandemic. Further work is required to understand future household behaviour changes in response to these price increases.

4.2 What is the future impact for NGED?

Over the course of ED2, NGED are interested in the implications for NGED's network planning and the implications for vulnerable customers connected to NGED's networks. On vulnerable customers, one of the conclusions from the previous section was that customers tended to change their electricity consumption behaviour less if they were vulnerable. We hypothesised a few reasons for this, but essentially the pandemic may not have instigated a change in working situation for vulnerable people, therefore they were less likely to change their behaviours. However a more likely prompt for behavioural change for this subset of customers is the cost of living crisis. In Figure 23, the final three months potentially illustrate this effect. Prices began to rise in winter 2021, with Ofgem's price cap moving up, reflecting the increase in gas prices at the time. This may be part of the explanation for financially poor customers reducing their consumption during this window. However, further work is required to understand this effect.

On the implications for NGED's network planning the above sub-section tells us that, in aggregate, there is likely to be little lasting impact of the pandemic on behaviours affecting electricity consumption; and any lasting impact may be affected by the ongoing cost of living crisis. However, as set out above, some specific groups (e.g. families with at least one person working from home) may show some persistence in their behaviour. So while the impact on the network is unlikely to be widespread in the future, where these customer groups are concentrated there may be particular geographic areas where the effect could impact NGED. As discussed in Section 3.4.1 NGED will consider network build in the context of how it thinks peak demand might develop at different substations. Our analysis in that section suggested there were a small number of constrained areas that might be at risk of being capacity constrained should future peak demand increase in a similar way as during the COVID-19 pandemic.

In this section we set out further analysis of the customers at these substations. Specifically, we investigate whether there are likely to be customers that persist in behaviours developed during the pandemic. We focus on the combination of two factors describe above: households that changed working patterns to work from home, and households with children.

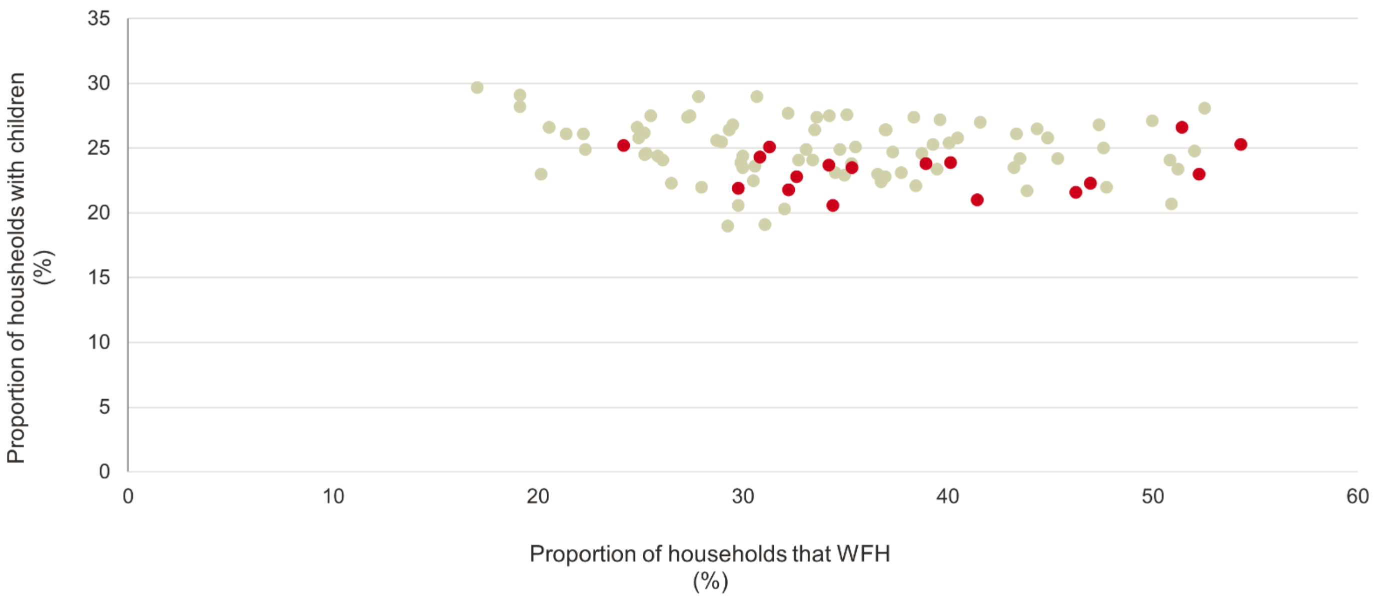
[48] As approximately 25% of the working population are parents that work from home, our sample underrepresents the population.(source: Labour Force Survey; Working Families Briefing)

We could not identify the characteristics of customers at individual substations, but using NGED's Network Capacity Map[49], we can identify the location of specific substations and map that to local authority districts. From there we are able to identify, by substation, the proportion of households with children and proportion of adults working from home in the substation's local authority. Figure 28 shows this for each substation which has limited demand headroom (identified by a red dot) and all other areas in NGED's network area (shown by the light green dots). Substations in the top right hand side of the chart are more likely to have a greater proportion of population that work from home and have children. Of the substations that are constrained, there are only a small minority that have high levels of working from home and high levels of children. The three substations with limited demand headroom and are located in local authorities with a high proportion of households which work from home and have children are:

- Kenilworth, in the Warwick local authority;
- Danesmoor, in the North East Derbyshire local authority; and
- Bruntingthorpe, in the Harborough local authority

These substations might require some additional reinforcement as a result of persistent impacts of the pandemic.

Figure 28 Proportion of households with children and proportion of adults working from home, by local authority



Source: NGED Network Capacity Map; ONS Homeworking in the UK, broken down by unitary and local authority districts, 2020 https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/labourproductivity/adhocs/13196homeworkingintheukbrokendownbyunitaryandlocalauthoritydistricts2020 ; ONS Census 2021

[49] https://www.nationalgrid.co.uk/our-network/network-capacity-map-application

The limitation of this analysis is that it does not tell us about other drivers that might influence more time spent at home. For example, there might be some areas where the commute to work is very short and therefore the inconvenience of returning to the office is less than in areas with a longer commute. Similarly, one local area might be dominated by an employer that has shut their office and encourages full time home-working. And, more generally, it does not acknowledge the large diversity in behaviours within an observable group of households (as described in Section 3.3.2).

5 Conclusions

The pandemic had a large impact on many aspects of life in Britain during 2020 and 2021. For many people, their daily routines were changed significantly, primarily in response to stay-at-home orders. There was significant difference in how individuals and households adapted to this in terms of the extent to which they had to adapt, and the costs associated with any change. This had significant impacts on domestic electricity consumption (in total and on its profile). These impacts varied significantly over time and across households.

The effect of the pandemic on electricity consumption was largest during the lockdown periods, and tended to affect households that previously spent days and evenings outside of the home (e.g. households where people previously went to work in offices, or households with children that were previously in school). For many other households, the impact of the pandemic was much lower and often appeared to have zero impact. And as the period went on, particularly during non-lockdown periods, the effect tended to diminish over time.

The habit formation framework gives an indication of the customer groups that are likely to persist their behaviour. The framework posits that individuals who receive a reward from changing their behaviour are more likely to persist in that new behaviour. Clearly the pandemic imposed a great individual cost on many people, as well as a significant social cost, so it is reasonable to assume that many behaviours would not persist unless necessary. However, we identified some areas where behaviour might persist: particularly the set of customers that work from home and receive some ongoing benefit from doing so (such as spending more time with other members of their household).

We also found that the impact of these behaviours persisting on NGED is unlikely to be large and systematic. For vulnerable customers in general there appeared to be no large increase in consumption, and therefore likely to be a small impact on bills, if any impact at all. For the network, the impact is likely to be targeted in specific geographic areas. This is because the customers that persist in their behaviour might be clustered in specific geographic areas, and the network areas that are more constrained are in specific locations. The overlap between network areas with a large proportion of customers that might persist in their behaviours and network areas with capacity constraints is small.

We acknowledge that this analysis should continue to be updated, particularly given the “cost of living crisis” associated with energy price inflation and the associated economic slowdown. This is likely to reduce household consumption. It will be important for the networks to identify what customer groups are likely to be affected, to what extent these groups decrease consumption and to understand interactions with the remaining impacts of the pandemic on consumption.

Annex 1 – Customer behaviour framework for habit formation

Elderly and vulnerable isolating

Application of the framework to working from home.

- **Trigger:** The trigger event is the pandemic, specifically the announcement on 23 March 2020 and the fact that elderly and clinically vulnerable people were advised to 'shield'.
- **Action:** This led to around 1.5m people in England shielding during March-August 2020, and January-April 2021[50], which meant a stricter set of self-isolation guidelines. Since then, a mass vaccination campaign has reduced the risk for these groups, but not eliminated it.
- **Reward:** Both rewards and costs are relevant to determining the net reward.
  - **Social reward** is the reward of staying at home, such as reducing the chance of passing COVID-19 to family or friends.
  - **Personal rewards** include not catching COVID-19, and making use of online/digital services (e.g. transitioning to online shopping).
- **Personal cost** is the cost of staying from home, such as loneliness.
- **Ease of persistence:** There are two ways this repeated action has primed continue working from home:
  - Individuals have adapted their lives to deal with increased isolation (e.g. ordering groceries online); and
  - Organisations have adapted to cater to those spending more time at home (e.g. virtual GP appointments, home catering, etc).

Based on this application of the habit formation framework, Figure 29 sets out our assessment of the impact of increased working from home on an individual basis.

[50] ONS, Coronavirus and clinically extremely vulnerable (CEV) people in England, https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronavirusandclinicallyextremelyvulnerablepeopleinengland/4aprilto23april2022

Figure 29 Increased working from home: individual impacts

Reward/cost from the action	Population group	Size of reward / cost	Chance of persistence
Not catching COVID	▪ 58% of over 70s were uncomfortable or very uncomfortable about leaving their homes in January. [1] ▪ This figure fell to 22% in August.	▪ Given the increased risk of hospitalisation and death relative to other age groups, the size of this reward is likely to be large, particularly pre-vaccine.	▪ Lower due to vaccine programme, but still some risk and prevalence for the elderly.
More convenient lifestyle from utilising online services	▪ Many elderly people remain digitally excluded, even after the pandemic – although there is some evidence that the pandemic increased the number of elderly customers online. [2]	▪ Rewards are likely to be greater for those who didn't previously use online services.	▪ Likely for those able to "get online", but less likely for the remaining customers that remain offline.
Lack of socialising with friends and family (i.e. loneliness)	▪ 50% of adults who live alone and are aged 65 and above had lockdown loneliness. [3] ▪ 14% of two adult households (at least one aged 65 and above) had lockdown loneliness.	▪ Greater cost for those who live alone. ▪ Greater cost for those who don't use social media/networks to stay in contact.	▪ Low chance of persistence for most given high vaccination rates.

Source: Frontier Economics

School-age children at home

Application of the framework to school-age children remaining at home more.

- **Trigger:** The main trigger was public health rules that required schools to shut, which led to children studying at home. Even with schools open, an additional trigger is the requirement for children to self-isolate due to having COVID-19 or being in close contact with someone with COVID-19.
- **Action:** Between May and June 2020, 87% of parents said a child in their household had been home-schooled because of COVID-19.[51] In June 2021, 11% of pupils (11-16) in state funded schools had COVID-19-related absence.[52]
- **Reward:** Both rewards and costs are relevant to determining the net reward.
  - **Social reward** is the reward of staying at home, such as reducing the chance of passing COVID-19 to family, friends, and staff in schools.
  - **Personal rewards** is the reward from studying at home, such as not catching COVID-19.
- **Personal cost** is the cost of studying at home, such as sub-optimal learning.
- **Ease of persistence:** There are two ways this repeated action has primed continue studying from home:
  - Pupils have obtained the necessary equipment to study from home (e.g. laptops) either from their school or self-bought; and

[51] ONS 2020, Coronavirus and homeschooling in Great Britain: April to June 2020

[52] Department of Education 2021, Attendance in education and early years settings during the coronavirus (COVID-19) pandemic, week 28 2021

- Schools have learned how to ‘manage’ virtual education.

Based on this application of the habit formation framework, Figure 30 sets out our assessment of the impact of increased studying at home on an individual basis.

Figure 30 Applying the habit formation framework to increased studying at home

Reward/cost from the action	Population group	Size of reward / cost	Chance of persistence
Not catching covid	▪ As of March 2021, 47% of parents with dependent children were “very or somewhat worried” about their child returning to school. **[1]**	▪ Low reward as pupils are less at risk compared to other age groups. ▪ Higher reward for risk averse and clinically vulnerable.	▪ Low given that children are generally less at risk compared to other age groups.
Poorer educational experience compared with in-person.	▪ There are 12 million people aged between 4 and 18 in Great Britain. **[2]**	▪ Bigger cost for pupils with less ability to buy necessary equipment to study from home. ▪ Also depends on whether schools provide the necessary equipment to study at home.	▪ Low given that schools are returning to in-person teaching. **[3]**
Central conclusion	▪ It seems reasonable to assume that, given the low risk to children versus older adults, there is less likely to be persisting risk-aversion, and remaining at home. ▪ Similarly, the learning experience is likely to be improved for all school-age children by attending school – therefore we might be able to conclude that this behaviour is unlikely to persist. ▪ A small increase in home schooling is likely to be insignificant in this context.		

Source: *Frontier Economics analysis*

Note: *1 – ONS 2021, 2 – ONS 2021, 3 - There has been an shift towards home schooling, this is ultimately very small (around 15,000 across the UK)*

Unemployed remaining at home

Application of the framework to unemployed people remaining at home more.

- **Trigger:** There are two potential triggers:
  - For the more risk averse, the initial spread of the virus triggered staying at home.
  - For the less risk averse, the government imposed stay at home guidelines.
- **Action:** The unemployment rate between March 2020 and June 2021 varied between 4.0% and 4.6%, with a peak of 5.2% in November 2020. See Figure 3 for more detail.
- **Reward:** Both rewards and costs are relevant to determining the net reward.
  - **Social reward** is the reward of staying at home, such as reducing the chance of passing COVID-19 to family or friends.
  - **Personal rewards** is the reward of being at home, such as not catching COVID-19, and spending less money away from home.
- **Personal cost** is the cost of being at home, such as loneliness.

- **Ease of persistence:** It is straightforward for the unemployed to remain at home given that many unemployment services, such as coaching, are available via phone call or video call.

Based on this application of the habit formation framework, Figure 31 sets out our assessment of the impact of increased staying at home on an individual basis.

Figure 31 Applying the habit formation framework to the unemployed spending more time at home

Reward/cost from the action	Population group	Size of reward / cost	Chance of persistence
Not catching covid	▪ As of September 2021, 17% of 16-69 year olds were "uncomfortable" or "very uncomfortable" about leaving home. **[1]**	▪ Large reward of staying home before vaccination programme, reducing over time given increasing vaccination rates. ▪ Higher for risk averse and vulnerable.	▪ Lower due to vaccine rollout, but still some risk for older &/or vulnerable workers.
Spending less money outside of home	▪ Between March 2020 and June 2021, the number of unemployed ranged between 1.4 million and 1.8 million. **[2]**	▪ This reward is likely to be relatively minor as the unemployed already have limited income and the savings from staying indoors are relatively small, e.g. not having to pay for travel as UC services are online.	▪ Some may like the online services whilst some may not, and the size of these groups are uncertain.
Loneliness	▪ Areas with a higher unemployment rate have higher proportions of residents who say they are often or always lonely. **[3]**	▪ Greater cost for those who are unemployed in a single person household compared to being unemployed in a multi person household.	▪ Higher persistence in multi-person households, lower persistence in single person households.
Central conclusion	▪ The overall picture is mixed with regards to the persistence of the unemployed remaining at home. ▪ The likelihood of persistence will depend on the trade-off between the convenience of using online services, and the cost of being lonely if the individual is in a single-person household. ▪ Persisting habits may impact a proportion of a reasonably small (<50%) population group.		

Source: Frontier Economics analysis

Note: 1 – ONS 2021, 2 – ONS 2021, 3 - ONS 2021

Furloughed employees remaining at home

Application of the framework to furloughed employees staying at home instead of working.

- **Trigger:** The announcement and implementation of the Coronavirus Job retention Scheme (CJRS) in March 2020.
- **Action:** Between July 2020 and June 2021 there have been at least 1 million employees on full-time furlough, with a peak of 4.5 million in 2020.[53] The furlough scheme ended on 30th September 2021.
- **Reward:** Both rewards and costs are relevant to determining the net reward.
  - **Social reward** is the reward of staying at home, such as reducing the chance of passing COVID-19 to family, friends, or colleagues.
  - **Personal rewards** is the reward of being on furlough, the chance to undertake informal labour, and not catching COVID-19.

[53] HMRC 2021, Coronavirus Job Retention Scheme statistics: 9 September 2021

■ **Personal cost** is the cost of being on furlough, such as not earning a full wage, and possible loss of skills and career development.

■ **Ease of persistence:** Once furlough was implemented, it was straightforward for furloughed employees to keep on receiving government contributions through the furlough scheme and stay at home.

Based on this application of the habit formation framework, Figure 32 sets out our assessment of the impact of increased staying at home for those furloughed on an individual basis.

Figure 32 Applying the habit formation framework to workers who were furloughed

Reward/cost from the action	Population group	Size of reward / cost	Chance of persistence
Not catching COVID	▪ As of Sept 2021, 17% of 16-69 year olds were "uncomfortable" or "very uncomfortable" about leaving home. **[1]**	▪ Large reward before vaccination programme, reducing over time given increasing vaccination rates. ▪ Higher for risk averse and vulnerable.	▪ Lower due to vaccine programme, but still some risk and prevalence for older workers.
Earning additional income through informal labour	▪ Unsurprisingly, there is little evidence regarding the number of furloughed employees who have engaged in informal labour whilst on furlough.	▪ Relatively greater reward for those on low incomes.	▪ Unless the furloughed job is part-time, it is unlikely that the individual can work formally and informally as furlough ends and so they return to their formal job.
Not receiving a full wage	▪ Between March 2020 and June 2021 there have been at least 2 million employees on furlough, with a government contribution of 80% for those dates. **[2]**	▪ Greater cost for those on lower incomes and those who financially support others, such as parents. ▪ Greater cost for those who don't undertake informal labour.	▪ End of furlough scheme means that the furloughed are likely to return to work, meaning that they will spend less time at home.
Loss of skills	▪ The furloughed would like to improve a greater number of skills when compared to the employed. **[3]**	▪ The size of the cost will depend on the person's desired career trajectory, and the career opportunities within their industry.	▪ End of furlough means that those who return to work can resume developing on-the-job skills.

Source: Frontier Economics analysis

Note: 1 – ONS 2021, 2 – ONS 2021, 3 - ONS 2021

Increased young people consolidating homes

Application of the framework to young people moving homes and consolidating places of residence.

■ **Trigger:** The trigger for young people moving home is the introduction of public health rules which required universities to stop in-person teaching.

■ **Action**: 42% is the proportion of 19 year olds who moved in with their parents (or parents-in-law), having lived independently of their parents prior to the COVID-19 outbreak.[54]

■ **Reward:** Both rewards and costs are relevant to determining the net reward.

  □ **Social reward** is the reward of returning home, such as reducing the chance of passing COVID-19 to family, friends, and university staff.

[54] *Zilanawala et al., 2020*. Household composition, couples' relationship quality, and social support during lockdown – Initial findings from the COVID-19 Survey in Five National Longitudinal Studies. London: UCL Centre for Longitudinal Studies.

- Personal rewards is the reward from studying at home, such as ease and saving money, and not catching COVID-19.

- **Personal cost** is the cost of studying at home, such as missing out on the social aspects of university.
- **Ease of persistence:** Following the initial action of moving home, which may have involved effectively moving out of the student accommodation, it is straightforward for the student to continue studying from, and living at, home.

Based on this application of the habit formation framework, Figure 33 sets out our assessment of the impact of young people consolidating their places of residence on an individual basis.

Figure 33 Applying the habit formation framework to young people returning to their family home

Reward/cost from the action	Population group	Size of reward / cost	Chance of persistence
Not catching covid	90% of students think that COVID presents a minor or major risk to their mental or physical health. [1]	Low reward as students are lower risk compared to older age groups. Higher reward for risk averse and clinically vulnerable.	Low given that students are low risk.
Ease of studying at home	Whilst there appears to be no data on ease, it's reasonable to assume that most students will find it easier to study at home than travel to university from their student home.	Reward is likely to be moderate across most students.	Low given that students are likely to prefer in-person teaching, which requires leaving home.
Saving money	It's reasonable to assume that most students will benefit financially from moving back home.	Greater reward if the student comes from a low income household. Students who can't work due to furloughed jobs may have a net cost.	Depends on net effect of saving money by living at home vs not being able to work.
Missing out on social aspects of university	55% of students have been dissatisfied with their social experience over the 2020-21 academic year. [2]	The cost is likely to be large across most students, in particular for those who are extroverted &/or less risk averse.	Low given that universities are returning to in-person teaching.

Source: Frontier Economics analysis
Note: 1 – ONS 2021, 2 – ONS 2021

Non-UK nationals leaving the UK

Application of the framework to non-UK nationals leaving the UK.

- **Trigger:** Given the timing of Brexit and the start of the pandemic, it is not clear which factor was the main cause of the increase in non-UK nationals leaving the UK.
- **Action**: 1.3 million non-UK born nationals left the UK between Q3 2019 and Q3 2020.[55]
- **Reward:** Both rewards and costs are relevant to determining the net reward.

[55] Economic Statistics Centre of Excellence, 2021

- **Personal rewards** are the rewards of leaving the UK, such as potentially being closer to some family and friends.

- **Personal cost** are the costs of leaving the UK, such as changing jobs and accommodation.
- **Ease of persistence:** For those who left because of Brexit, the initial action of leaving the UK will have required a large amount of effort in terms of changing jobs and finding new accommodation, and so there will be little effort required to continue living in the new location. For those who left due to the pandemic, it is likely that most plan to return to the UK to the same job and accommodation, meaning that long-term persistence will be difficult.

Based on this application of the habit formation framework, Figure 34 sets out our assessment of the impact of non-UK nationals leaving the UK on an individual basis.

Figure 34 Applying the habit formation framework to non-UK nationals leaving the UK

Reward/cost from the action	Population group	Size of reward / cost	Chance of persistence
Being closer to family and friends	It is reasonable to assume that most of the 1.3 million will consider moving to a location where they are close to friend and family.	Reward is likely to be greater for those whose closest friends and family live outside of the UK.	Easier to travel to see friends and family in the future, so some may return to the UK.
Changing jobs and accommodation	Those who left because of the pandemic are likely to keep their same job and seek temporary accommodation. Those who left because of Brexit are likely to switch jobs and seek permanent accommodation.	Cost is likely to be greater for those who have left due to Brexit, and the cost is smaller for those who are temporarily relocating.	Medium (regardless of trigger) as the time and effort (i.e. cost) to change accommodation and jobs will be short-lived.
Central conclusion	The chance of persistence depends strongly on the trigger. Those who moved abroad because of Brexit are more likely to be moving permanently (i.e. high persistence), whilst those who moved abroad because of the pandemic are less likely to be moving permanently, and at some point plan to return to the UK to work. These workers therefore have jobs that are suitable for WFH, and they may WFH to some extent upon returning to the UK.		

Source: Frontier Economics analysis

Note: 1 – ONS 2021, 2 – ONS 2021

Annex 2 - Methodology for estimating the counterfactual

Dr Ellen Zapata-Webborn and Dr Eoghan McKenna

University College London, 2022

Introduction

Here we describe the predictive modelling used to estimate what electricity consumption would have been without the influence of the COVID-19 pandemic. We begin by describing the datasets used including estimates of the sample representativeness. Then we describe the counterfactual modelling process including models used, model selection and evaluation, and finally potential weaknesses and considerations for interpreting the results.

Datasets

The Smart Energy Research Lab (SERL)[56] has been recruiting households in Great Britain from 2019 to 2021 in order to collect their half-hourly electricity and gas smart meter data for research. Data collection[57] is ongoing with data for some of the participants recruited earliest dating back to August 2018. When signing up participants were requested to complete a survey about their home and household, and in May 2020 the first wave of recruits was sent a survey about the experiences during the first COVID-19 lockdown. In addition to the smart meter and survey data, hourly reanalysis weather data from ECMWF[58] and Energy Performance Certificate (EPC) data are linked at the household level, along with region and index of multiple deprivation (IMD) quintile.

These datasets are available to accredited researchers via a secure-lab environment following a project application for research in the public interest. Project require university ethics approval, approval by the SERL Data Governance Board, and accredited researcher status for all those with data access. For more information about how to access SERL data visit the SERL website.

Data filtering

In every 'required month'[59] a participant was required to have at least 50% of the days in the month with no missing or invalid[60] half-hourly reads in order to be included in the sample. While a full year of historic data would have been preferable, due to the timing of the SERL recruitment waves, requiring data further back than April 2019 significantly reduced the sample size and the gains from additional historic data were not

---

[56] www.serl.ac.uk

[57] Elam, S., Webborn, E., McKenna, E., Oreszczyn, T., Anderson, B., Few, J., Pullinger, M., European Centre for Medium-Range Weather Forecasts, Ministry of Housing, Communities and Local Government, Royal Mail Group Limited. (2022). Smart Energy Research Lab Observatory Data, 2019-2021: Secure Access. [data collection]. 5th Edition. UK Data Service. SN: 8666, DOI: 10.5255/UKDA-SN-8666-5.

[58] https://www.ecmwf.int/

[59] 'Required months' are April 2019 – February 2020 (historic data for model training) and April – May 2020 (approximately the first UK COVID-19 lockdown).

[60] The data has gone through initial processing by the SERL team and suspicious or clearly invalid reads have been flagged.

worth the sample size reduction. For households that did have more historic data than the minimum requirement this data was included for model training. Similarly, households were not required to have data beyond May 2020 as this covered the first national lockdown. Data collection for a household stops when they withdraw consent to participate or move house. Following participant filtering 586 participants remained in the 'COVID-19 analysis sample'.

Smart meter data was filtered to remove any days without a full set of valid half-hourly reads. March 2020 data was also removed as some households started reacting to the pandemic at different times during the month, so it was unsuitable for model training or use in predictions.

Data preparation

Daily periods

Each day was split into eight sub-periods (which we refer to as 'periods') because aggregation improves model predictive power and half-hourly predictions were not required for the analysis. The periods were defined by considering periods of fairly static demand in daily profiles from the SERL Annual Report

Table 5: Definition of daily periods.

Period	Description	Time span (local time)	Duration (hours)
A	Night-time	23:00 – 06:00	7.0
B	Early morning	06:00 – 08:30	2.5
C	Mid/late morning	08:30 – 12:00	3.5
D	Lunchtime	12:00 – 14:00	2.0
E	Mid-afternoon	14:00 – 16:00	2.0
F	Late afternoon	16:00 – 17:30	1.5
G	Evening peak	17:30 – 20:00	2.5
H	Mid-late evening	20:00 – 23:00	3.0

Smart meter data

Following data filtering a number of pre-processing steps were taken to prepare the data for model training and testing and for analysis.

1. Each half-hourly read was linked to its period. Because period A starts at 11pm, an 'assigned day' variable was created so that rather than a day starting at midnight, the assigned day began at 11pm the night before (in local time). All analysis was done using assigned days so that periods did not span multiple days and henceforth 'day' refers to assigned day unless stated otherwise.

2. Additional variables were added to each half-hourly electricity read:
    a. Period (see above)
    b. Weekend/weekday flag (true if the day is on a weekend)
    c. Public holiday flag (true if a public holiday in England)
    d. Month
    e. Year
    f. Sinusoidal calendar variables (see below)

We add two variables to capture the cyclical nature of the year (i.e. that 31st December and 1st January are approximately the same rather than polar opposites):

$$\text{sin\_day} = \frac{1}{2}\sin\left(\frac{2\pi d}{D}\right) + \frac{1}{2}$$
$$\text{cos\_day} = \frac{1}{2}\cos\left(\frac{2\pi d)}{D}\right) + \frac{1}{2}$$

where d is the position of the day in the year, *i.e.* 1st Jan has d = 1 and D is the total number of days in the year (364 or 365). We normalised the sinusoidal transformations (restricted to between 0 and 1) for the benefit of the neural network models.

Weather data

The ECMWF weather data has no missing values. Linear interpolation was used to impute half-hourly values as the original data is hourly. Three variables were selected for use in the models:

- Temperature (instantaneous on the hour) at 2m above surface level
    - Originally in °K, converted to °C
    - *2m_temperature_K* in the SERL dataset
- Solar radiation (cumulative from the preceding hour)
    - Measured in Jm^{-2}
    - *surface_solar_radiation_downwards* in the SERL dataset
- Total precipitation (cumulative from the preceding hour)
    - Measured in m
    - *total_precipitation* in the SERL dataset.

The electricity consumption reads are half-hourly and cumulative from the previous half-hour. To impute the temperature data an approximation for instantaneous temperature at 15 minutes past the hour and 15 minutes to the hour was the best match for the smart meter data. The following formulas were used to interpolate temperature at these times, where *T* is the original temperature data, $\hat{T}$ is interpolated temperature, *hh:00* is to match a smart meter reading on hour *hh*, and *hh:30* is to match with a smart meter reading at half past hour *hh.*

$$\hat{T}(hh{:}00) = 0.75 * T(hh{:}00) + 0.25 * T(hh - 1{:}00)$$

$$\hat{T}(hh{:}30) = 0.75 * T(hh{:}00) + 0.25 * T(hh + 1{:}00)$$

Cumulative variables solar radiation and total precipitation were interpolated as follows, where *V* is the cumulative variable and $\hat{V}$ is the interpolated variable.

$$\hat{V}(hh{:}00) = \frac{1}{2}\big(V(hh{:}00)\big)$$

$$\hat{V}(hh{:}30) = \frac{1}{2}\big(V(hh + 1{:}00)\big)$$

Sample representativeness

To understand how our results may generalise, it is useful to consider how our 'COVID-19 analysis sample' (the 586 households in this study) compare with the full 'SERL sample' of households recruited to the SERL project and, where possible, with the population of England and Wales. To do this we use information about a household's location, answers to the SERL survey (optionally) completed upon sign up, and a COVID-19 survey sent in May 2020. 575 households in the COVID-19 analysis sample (98%) answered the SERL survey either fully or in part; 367 (63%) for the COVID-19 survey.

Table 6 shows how the COVID-19 sample compares with national estimates and the full SERL sample. Note that the sample is broadly representative regionally, except that the COVID-19 sample does not represent households in the North of England or in Scotland due to no participants having been recruited from those regions in time for sufficient historic data to be included in our sample. The figures that follow show how the COVID-19 analysis sample compares with the SERL sample (~13,000 households) for all households with the relevant data.

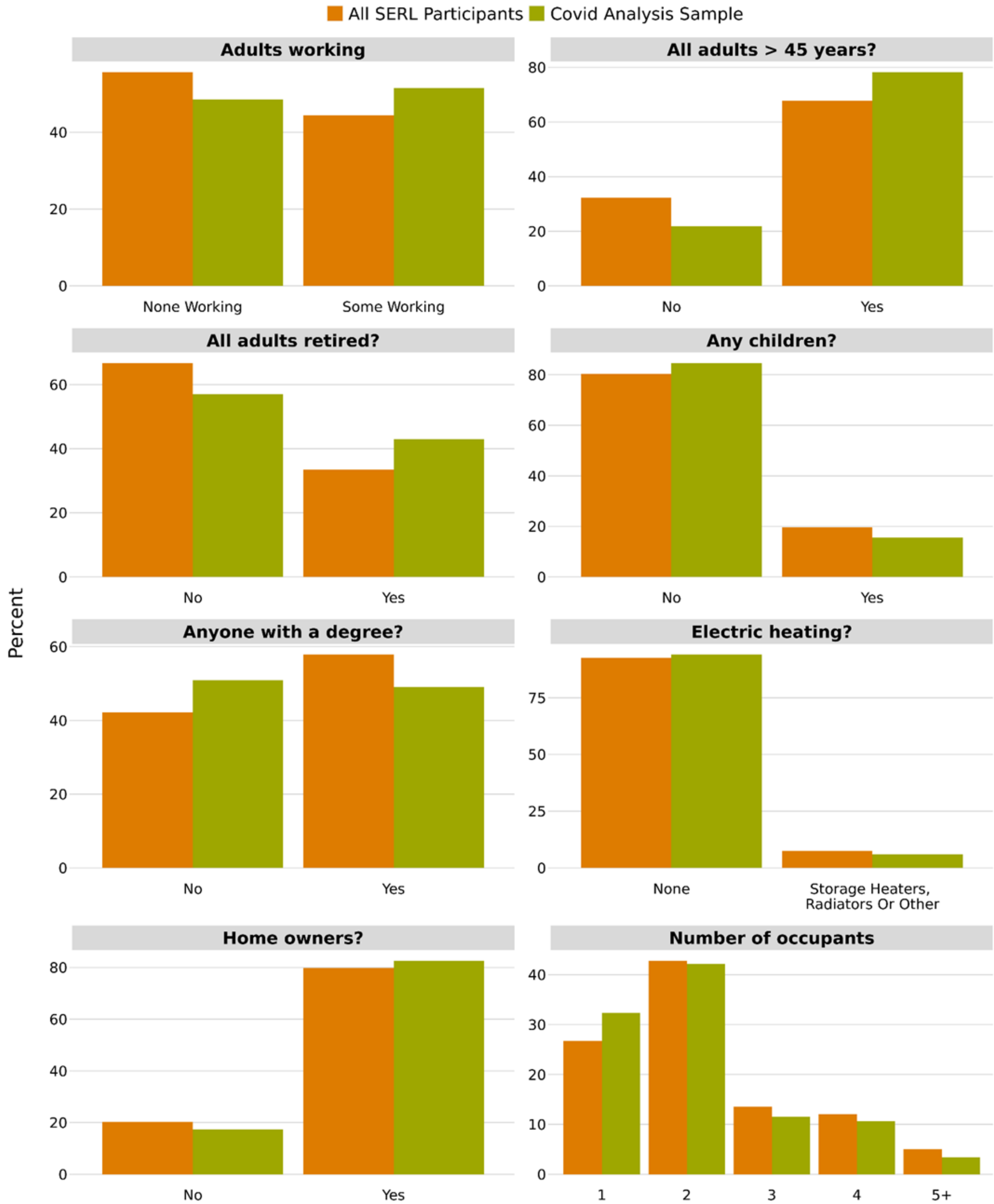
While we consider it beneficial to present a study of the representativeness of our sample, it is not necessary for the analysis that our sample is perfectly representative, as we do not suggest that the impact of COVID-19 on our sample is reflective of the impact throughout England and Wales. Our model is used to identify different patterns of response and how they correlate to household characteristics, from which we can draw potential wider implications.

Table 6: Representativeness of the COVID-19 sample compared with national estimates and the full SERL sample/those who answered the SERL COVID-19 survey. Blank cells indicate no data available.

Attributes	**COVID-19 sample compared to…**		
	National Estimate	**SERL sample**	**SERL COVID-19 survey sample**
Affluent areas	Over-represents	Over-represents	
Deprived areas	Under-represents	Under-represents	
Those reporting they are not struggling financially	Over-represents	Under-represents	

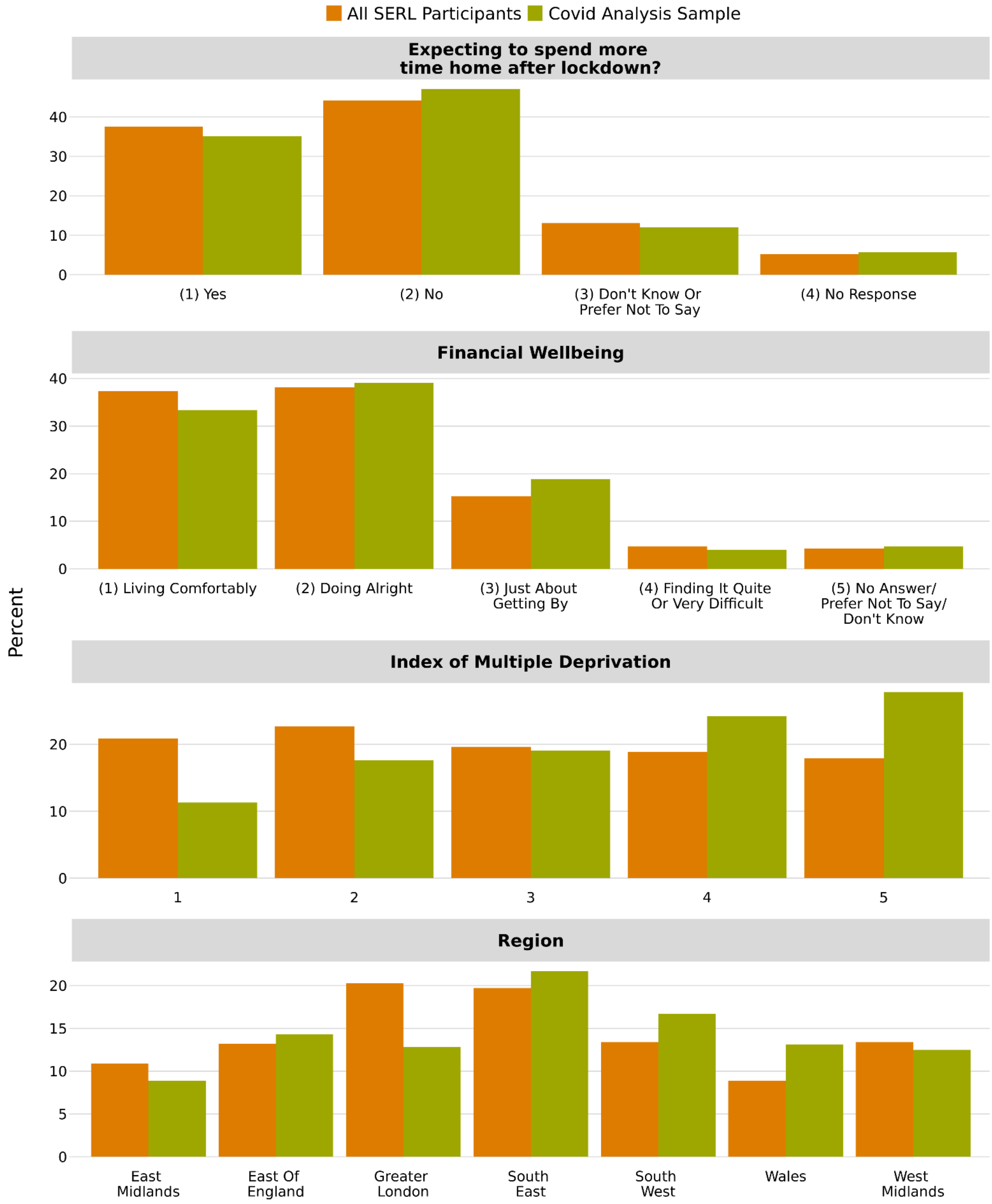
Those reporting they are struggling financially	Under-represents	Over-represents	
Those not expecting to spend more time at home after lockdown			Slightly over-represents
Homes with electric heating		Representative	
Single-occupant households	Over-represents	Over-represents	
Large households	Under-represents	Representative	
Retired households		Over-represents	
Owner occupiers	Over-represents	Over-represents	
Childless households		Slightly over-represents	
No occupants with a degree		Over-represents	

Figure 35 Comparison of the COVID-19 analysis sample and the wider SERL sample in terms of representativeness



Source: Frontier Economics / UCL analysis using SERL data

Figure 36 Comparison of the COVID-19 analysis sample and the wider SERL sample in terms of representativeness



Source: Frontier Economics / UCL analysis using SERL data

Counterfactual modelling

Introduction

What impact did COVID-19 and the resulting lockdowns have on electricity consumption in households in England and Wales? To answer this question, we need a way to estimate what would have happened if there had been no pandemic. We call the estimate for what would have happened the 'baseline' or 'counterfactual'.

Other studies have attempted to determine the impact of COVID-19 on domestic electricity demand[61]. Published studies have focused on the impact of COVID-19 during 2020, rather than the continued impact into 2021, and common approaches to establishing the counterfactual have been to use the same period in the previous year, or (when considering the first lockdown only) compare lockdown demand with demand in the few weeks before lockdown started. These approaches are severely limited because the weather (particularly in the first national lockdown) was very different from any weather in the previous year, and the effects of COVID-19 were already being felt in the few weeks before the national lockdown started. They also prevent any long-term monitoring of consumption changes. Some studies have used predictive modelling (which we do here), such as a Cornwall study of 50 homes[62] which used mixed linear regression with internal temperature data. However the sample size was small and mostly long-term sick, disabled or retired residents so their findings were largely not statistically significant. Studies of national-level electricity demand (including industrial demand) have shown a reduction in electricity demand caused by the lack of industry, such as Mehlig *et al.*[63] which created a counterfactual model using ordinary least squares regression.

To determine the baseline for each household we developed two types of models commonly used for predicting domestic electricity demand. These models use data from before the pandemic to learn the relationship between electricity consumption and the calendar and weather variables. Each household is modelled separately to allow for different relationships between the independent and dependent variables. Different variations of the models were compared to find the one with the lowest error for each household, so not only does each household have a different model, but they have different types of model according to which worked best for each household. Two types of model were tested (including variations of each): elastic net regression and neural networks. The model outputs are a set of predictions from April 2020 to the end of December 2021 for each household for each period of the day (Table 5) alongside the observed electricity consumption during this time. The units used were mean hourly electricity consumption in Wh so that the periods were comparable despite being of varying length.

---

[61] For a review see, for example, Krarti, M. and Aldubyan, M. (2021). Review analysis of COVID-19 impact on electricity demand for residential buildings. *Renewable and Sustainable Energy Reviews*. **143**

[62] Meneer, D. *et al.* (2021). Changes in Domestic Energy and Water Usage during the UK Covid-19 Lockdown Using High-Resolution Temporal Data. *International Journal of Environmental Research and Public Health*. **18**(13).

[63] Mehlig, D. *et al.* (2021). The Impact of the UK's COVID-19 Lockdowns on Energy Demand and Emissions. *Environmental Research Letters*. **16**(5).

Elastic net regression

Introduction

Regression analysis seeks a functional relationship between response variable(s) (in our case electricity consumption) and predictor/independent variables (such as weather, day of the week). Linear regression is a of regression that assumes a linear relationship between predictor and response variables. Linear regression models are the simplest family of predictive models and are therefore a good place to start. Challenges for regression include 'over-fitting'[64] and lack of model. Using a regularisation (or 'shrinkage') method with regression is a way to penalise large coefficients and/or the inclusion of many predictor variables, which can help address these issues. Regularisation reduces model variance at the expense of a small increase in model bias. Three methods are available: ridge regression which penalises large coefficients (but not to zero), lasso regression which penalises a large number of predictor variables, and elastic net regression[65]; a combination of the two. Elastic net regression has previously been found to outperform other methods for electricity consumption prediction[66] and so this was the first model we developed and tested. In elastic net regression, for n response data points y_i, and p predictor variables x_j, the coefficients β_j for each predictor the are determined by minimising the residual sum-of-squares plus the elastic net penalty:

$$\sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij}\right)^2 + \lambda\sum_{j=1}^{p}\left(\alpha\beta_j^2 + (1-\alpha)|\beta_j|\right)$$

Tuning parameters $\alpha \in [0,1]$ and $\lambda \geq 0$ are determined by trialling different values to determine the combination which gives the optimal performance. $\alpha = 0$ is the ridge penalty; $\alpha = 1$ is the lasso penalty, and in between is a combination of the two. Prior to minimisation the response is centred (y_i sum to 0) and the predictors standardised (for each predictor variable j, $\sum_i x_{ij} = 0$ and $\sum_i x_{ij}^2 = 1$).

We performed elastic net regression using R packages glmnet[67] and caret[68] with 10-fold cross-validation and tune length 10.

Predictor variables and regression formula

We found that most elastic net regression models gave better performance when each period was modelled separately. In this case the regression formula takes the form

---

[64] Overfitting is when a model fails to generalise to new data because it has become too closely aligned to training data and missed the general trends.

[65] Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *J. R. Statist. Soc. B*. **67**(2), pp301-320.

[66] Sâtre-Meloy, A. (2019). Investigating structural and occupant drivers of annual residential electricity consumption using regularization in regression models, *Energy*, **174** p148-168.

[67] Friedman, J., Hastie, T., Tibshirani R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of Statistical Software, 33(1), 1-22. https://www.jstatsoft.org/v33/i01/

[68] Kuhn, M. (2021). caret: Classification and Regression Training. R package version 6.0-90. https://CRAN.R-project.org/package=caret

$$\boldsymbol{y} = \beta_0 + \sum_{j=1}^{p} \beta_j \boldsymbol{x_j}$$

where bold font represents a vector of length n and predictors $\boldsymbol{x_j}$ are:

- Mean temperature
  - Over the period
  - Over the period * weekend interaction term[69]
  - Over the day
  - Over each of the 3 previous days (3 separate variables)
- Mean solar radiation
  - Over the period
  - Over the period * weekend (True/False) interaction term
  - Over the day
  - Over each of the 3 previous days (3 separate variables)
- Mean total precipitation
  - Over the period
  - Over the period * weekend (True/False) interaction term
  - Over the day

For some households the model performed better when considering all periods simultaneously, in which case the period variable was included in the model formula, and as an interaction term before all other variables currently with no interaction.

Recall that the elastic net regression eliminates inconsequential variables by penalising an excess number of predictor variables, therefore the optimized models may not include all of the above. Temperature and solar radiation lead to heat gains in buildings which can affect electricity use on later days (mainly if electric heating exists) so the model has the potential to use these weather variables from the past. In contrast, precipitation from a few days' previous is unlikely to affect current consumption. The interaction with the weekend variable allows a household to respond differently to weather conditions if it is a weekend. For example, you may go to work on weekdays irrespective of the cold, but at the weekend you might choose to stay inside and keep warm. Or when it is hot to do the washing so it can dry outside, but only if you're home (more likely at the weekend). Each household has its own model so they can have different variables and relationships between the variables from one another.

Initially all periods were considered together, but period A showed much higher bias. Since A is the night-time period, it makes sense that consumption at night may need a very different model to consumption during the day, when the weather has little impact (particularly as that day's weather is yet to occur). So period A was modelled separately, and then all periods modelled separately, and the most accurate model selected for each household.

[69] The 'weekend' variable is true if the day is Saturday or Sunday, otherwise false.

Performing elastic net regression in R

Elastic net regression was performed using the caret[70] and glmnet[71] packages in R. 10 values for each of tuning parameters α and λ were tested using the caret function 'train' and 10-fold cross-validation. The best-performing tuning parameters were then selected to train the predictive model using the full set of training data (all available data before the start of March 2020).

Neural network

Introduction

Similar to the elastic net regression model described above, a neural network (or 'artificial neural network') is a type of model that is used to estimate a function that describes the relationship between a set of input variables and one or more output variables. In this case the inputs are calendar and weather variables, and the output is electricity consumption for an individual household.

A neural network consists of a network of multiple 'neurons' or 'units' which are connected in 'hidden' layers so that the outputs of the units in one layer are used as inputs to the units in the next layer. This is known as a 'feedforward neural network'. Each unit is represented by an equation similar to the elastic net regression equation: a linear sum of the inputs multiplied by parameters known as weights and a bias term. The output of this equation is then passed through an 'activation function' which is usually a non-linear function. A neural network can consist of many such units and the benefit of this is that they can estimate highly non-linear relationships between input and output variables. This is a strength of neural networks compared to a simpler model such as the elastic net regression model described previously and a reason why we choose this type of model to complement it.

While there is an extensive literature on the use of neural networks to model energy consumption in buildings with several reviews available[72] [73] [74] [75] the majority are focussed on *forecasting* applications, and there are few studies that use neural networks to predict a high-resolution (e.g. hourly) energy consumption baseline counterfactual for residential buildings. Yalcintas[76] trains a neural network with a single hidden layer and using calendar and weather variables as inputs to predict hourly energy consumption for a hotel and estimate energy savings for retrofit energy efficiency measures. The ASHRAE Great Energy Predictor

---

[70] Kuhn, M (2021). Caret: Classification and Regression Training. https://CRAN.R-project.org/package=caret

[71] Friedman, J., Hastie, T., Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of Statistical Software, 33(1), 1-22. https://www.jstatsoft.org/v33/i01/

[72] Foucquier, S. Robert, F. Suard, L. Stéphan, A. Jay, (2013). State of the art in building modelling and energy performances prediction: A review, Renew. Sustain. Energy Rev. 23 272–288. https://doi.org/10.1016/j.rser.2013.03.004.

[73] A.T. Nguyen, S. Reiter, P. Rigo, (2014). A review on simulation-based optimization methods applied to building performance analysis, Appl. Energy. 113 1043–1058. https://doi.org/10.1016/j.apenergy.2013.08.061.

[74] K. Amasyali, N.M. El-Gohary, (2018). A review of data-driven building energy consumption prediction studies, Renew. Sustain. Energy Rev. 81, 1192–1205. https://doi.org/10.1016/j.rser.2017.04.095.

[75] Y. Wei, X. Zhang, Y. Shi, L. Xia, S. Pan, J. Wu, M. Han, X. Zhao, (2018). A review of data-driven approaches for prediction and classification of building energy consumption, Renew. Sustain. Energy Rev. 82 1027–1047. https://doi.org/10.1016/j.rser.2017.09.108

[76] Yalcintas, M. (2008). "Energy-savings predictions for building-equipment retrofits." Energy and Buildings 40.12: 2111-2120.

III competition[77] focussed on predicting hourly baseline energy consumption for non-domestic buildings and several of the winning entries used neural networks as part of an 'ensemble' approach – this consists of training many different models and creating a final model which learns how to combine these together to create an output that is more accurate than any one individual model. Here we employ a simple type of ensemble approach, by training two types of model (elastic net regression and neural network) and selecting the best performing one for each household.

Source data

The 'intermediary' files used in the elastic net regression model. One 'training' file and one 'prediction' file for each participant containing observations for the defined dates for these periods.

Pre-processing

Observations removed where demand was less than or equal to zero. Weather and active electricity import variables were normalised by subtracting the mean and dividing by the standard deviation of the relevant variable as calculated during the training period.

Sine and cosine transformations of calendar variables were of the form:

$$\text{sine\_month} = \sin\left(\frac{2\pi m}{M}\right)$$

$$\text{cosine\_month} = \cos\left(\frac{2\pi m)}{M}\right)$$

where *m* is the position of the month in the year, i.e. January has *m* = 1 and *M* is the total number of months in the year (12).

Model inputs and outputs

Inputs:

- Calendar variables
  - Half-hour of the day: dummy variables (leave one out) and sine and cosine transformation
  - Day of the week: dummy variables (leave one out) and sine and cosine transformation
  - Month of the year: dummy variables (leave one out) and sine and cosine transformation
- Weather variables
  - Normalised surface solar radiation downwards (J m-2)
    - For the half hour
    - Mean of the day
  - Normalised 2m temperature (°K)
    - For the half hour
    - Mean of the day
  - Normalised Total precipitation (m)
    - For the half hour

[77] Fontanini & Jeff Haberl (2020). The ASHRAE Great Energy Predictor III competition: Overview and results, Science and Technology for the Built Environment, 26:10, 1427-1447, DOI:
10.1080/23744731.2020.1795514.

- Mean of the day

Month of the year variables were not included if a participant did not have data for each month of the year.

Output: Normalised active electricity import (Wh).

Cross-validation

10-fold cross-validation was used to calculate the trained models out-of-sample error.

Model hyper-parameters

The structure of the neural networks consisted of two hidden layers each with eight units, and an output layer with a single unit. Layers were fully connected. Sigmoid activation functions were used for both hidden layers, and a linear activation function for the output layer. Regularisation was not used.

Training

The cost function used was the mean squared error of the model output compared to the target observations.

The cost function was minimised using an optimiser implementing a BFGS with Jacobian method and a maximum number of iterations of 500.

For each participant a separate model was trained for each cross-validation fold to allow estimation of the out-of-sample error.

Prediction

For each participant the cross-validation predictions are aggregated to produce a single file for the training period which contains the out-of-sample predictions for the 10 cross-validation models. This file is used to estimate the out-of-sample error / model performance for each participant.

For each participant, then a final single model using the same hyperparameters is trained on the full training data. This final model is then used to create counterfactual predictions during the prediction period using input data from the prediction period. These final predictions can be compared to the actual observations during the prediction period to estimate the change in electricity demand associated with the COVID-19 pandemic.

Post-processing

To allow the performance of the neural network models to be compared with the elastic net regression and to reduce model error, the final predictions for each participant are aggregated by calculating the mean value for each month of the year and defined sub-daily period.

Model performance

Model performance is estimated by calculating the error between the cross-validation out-of-sample predictions and the actual observations. Two error types are calculated: the normalised mean bias error (NMBE), and the coefficient of variation root mean squared error (CVRMSE).

Hardware and software

All analysis was performed in the UKDS Secure Lab AWS environment. Programming was implemented using python 3.9.12 in the spyder IDE 5.1.5 and using pandas 1.4.2, numpy 1.22.3. Gradient descent optimisation was performed using scipy 1.7.3.

Model selection and evaluation

For each household a number of models were tried to determine which was most suited for electricity use prediction. Model parameters were tuned using 10-fold cross-validation and the optimal parameters selected for predictions. Pre-pandemic data was used for training and testing which gives estimates for the error and bias of each model with the selected parameters. To compare, for example, a neural network model with an elastic net regression model we can compare the error and bias of these models from the cross-validation process. Tweaks could be made to the model structure/formula to improve the model, and the effects evaluated using these error and bias estimates.

Error is a measure of how accurate the model is in predicting the observed values; the smaller the error, the greater the accuracy. Bias is a measure of systematic over- or under-prediction of the observed values. During model selection and evaluation both error and bias were considered, as model accuracy is important, but an accurate model that, say, over-predicts demand at certain times of the day or months of the year could lead to results that appear to be significant that are not, or trends hidden by systematic bias in the predictions. Ultimately while normalised mean bias error (NMBE) was considered (particularly when exploring ways to improve the models), the choice of model for each household was the model with the lowest coefficient of variation of the root mean squared error: CV(RMSE). Our bias and error metrics are defined as follows:

$$\mathrm{NMBE} = \frac{1}{\mathrm{n}-1}\,\frac{\sum(\hat{y}_{\mathrm{i}} - \mathrm{y}_{\mathrm{i}})}{\bar{\mathrm{y}}}$$

$$\mathrm{CV(RMSE)} = \frac{1}{\bar{y}}\sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{n-1}}$$

Where n is the number of observations, y_{i} is the i^{th} observation, $\hat{y}_{\mathrm{i}}$ is the i^{th} prediction, and $\bar{y}$ is the mean of the observations. NMBE > 0 means the model tends to over-predict.

The following four model choices were compared to select the best model for each household (number in brackets shows the number of households for which this was the optimal model choice):

- Elastic net regression
  - All periods used for training and testing *(0 households)*

  - Period A modelled separately from the other periods *(21 households)*
  - Each period modelled separately *(42 households)*
- Neural network *(523 households)*

Variations on the elastic net regression were considered because we found that period A had much greater bias than the other periods. Period A is night-time demand, and so it is very likely that demand during period A is not influenced by the predictor variables such as weather or day of the week in the same way as day-time periods. Thus relationships determined by day-time periods are negatively influencing the predictive power of the model for night-time demand. For some households modelling all periods separately was beneficial, though mostly it made little difference.

Final model error and bias

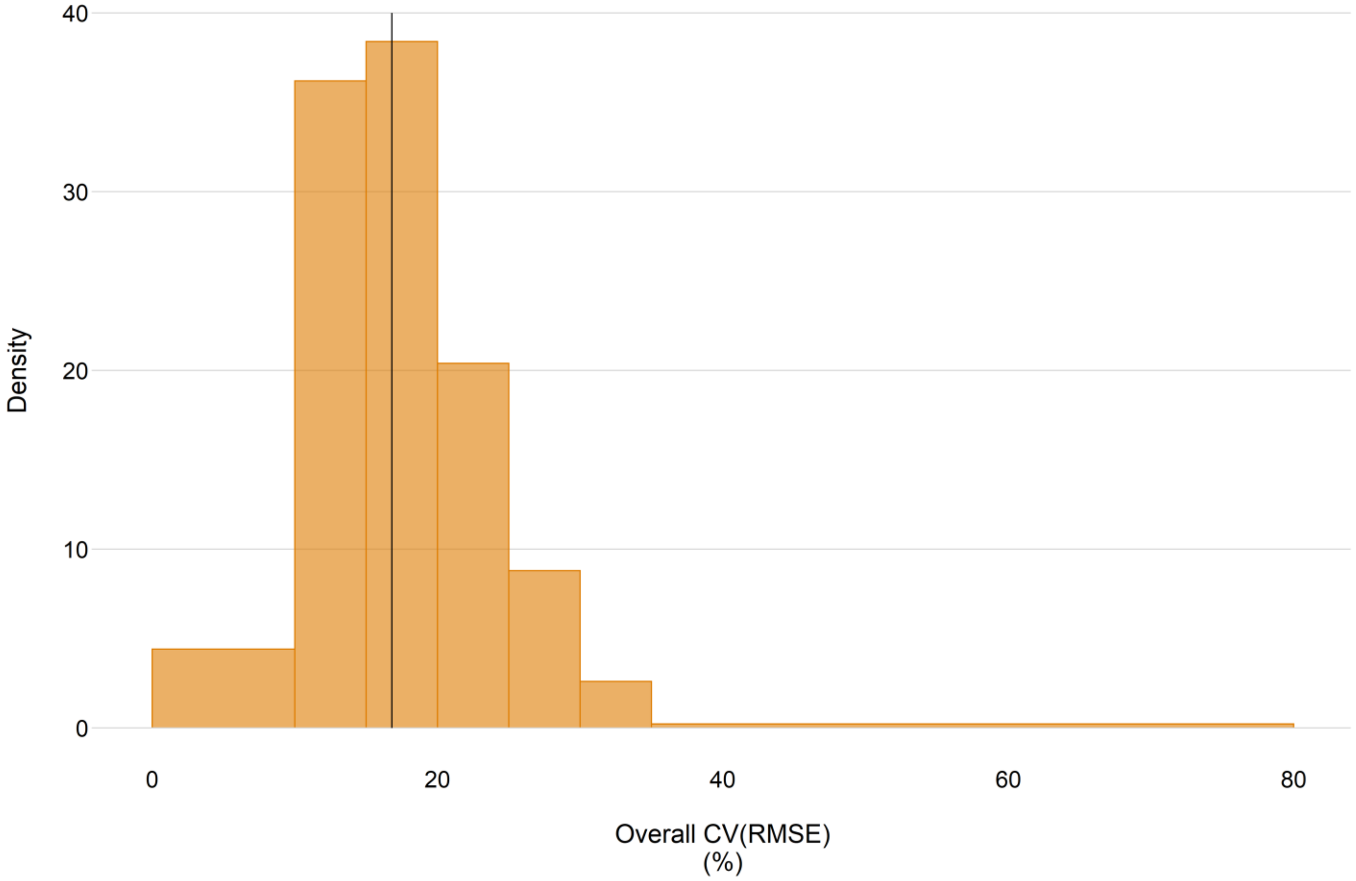
In this section we explore the error and bias of the final household models overall, by month and by period of the day.

Overall error and bias

Overall the error for the households in our sample is between approximately 0 and 80%, with a median of 16.8%. Note that for statistical disclosure control (SDC) we have combined histogram bins so that each bin represents at least 10 households, in line with SERL's SDC policy. We also use approximations for medians throughout this analysis by taking the average of at least 10 households with a value very close to the median, so that the value for an individual household is not revealed. Typical criteria for model selection are to require CV(RMSE) < 30% or 20% at an hourly resolution and CV(RMSE) < 15% at a monthly level[78]. Since our periods are much closer to hourly than monthly, most households have error lower than standard criteria, and overall error is significantly lower than the 30% standard.

[78] Ruiz, G. and Bandera, C. (2017). Validation of calibrated energy models: Common errors. *Energies* **10**(10).

Figure 37 Overall error for the COVID-19 sample

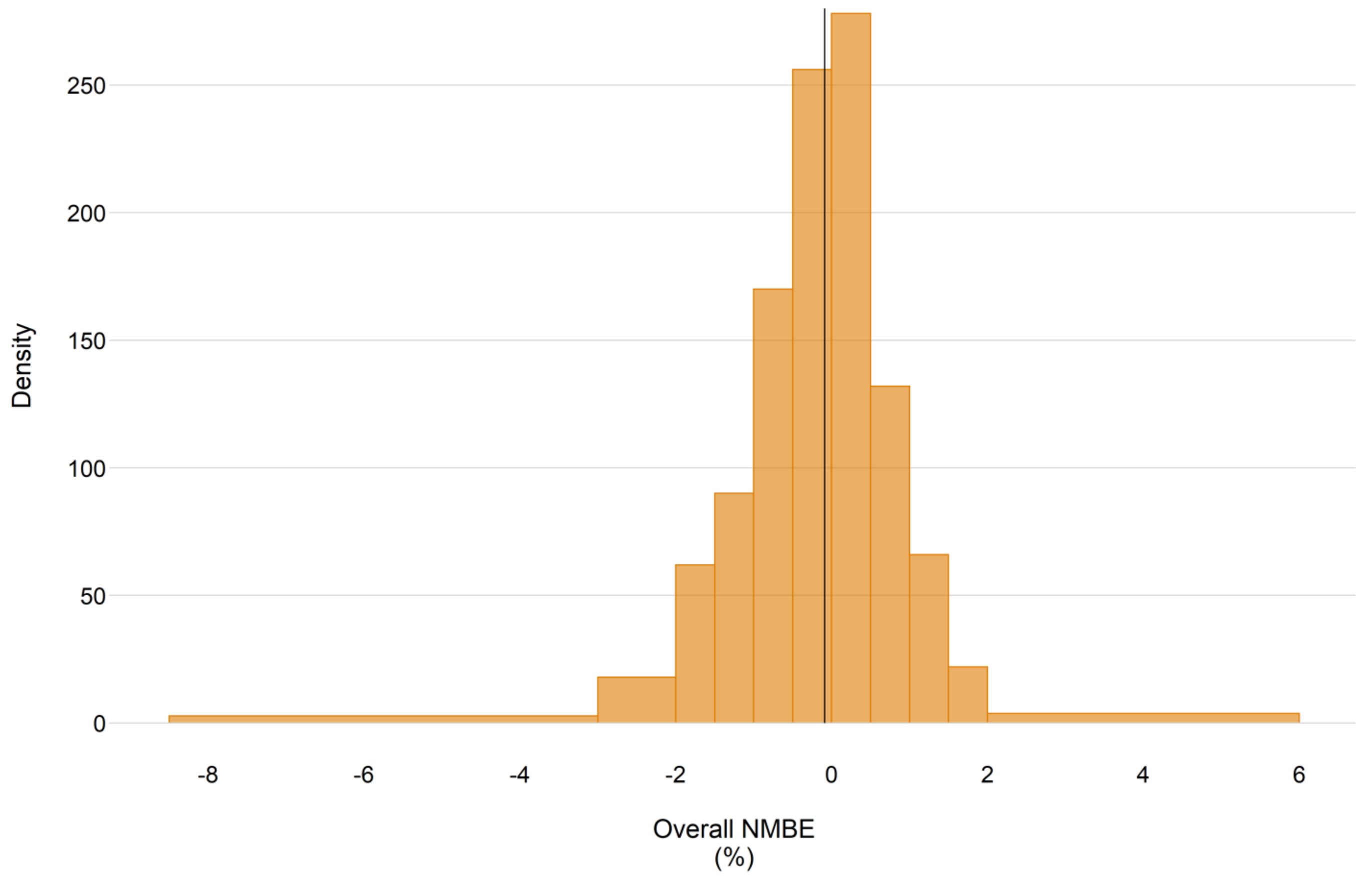


Source: Frontier Economics / UCL analysis using SERL data

Note: Black line indicates an approximation of the median; 16.8%.

The same source recommends NMBE within ±5% or 10% at an hourly resolution (NMBE within ±5% or 20% depending on the standard at a monthly resolution). The figure below shows overall bias well within these ranges for almost households, and overall (median) bias extremely close to zero at less than 0.09%.

Figure 38 Overall model bias for the population

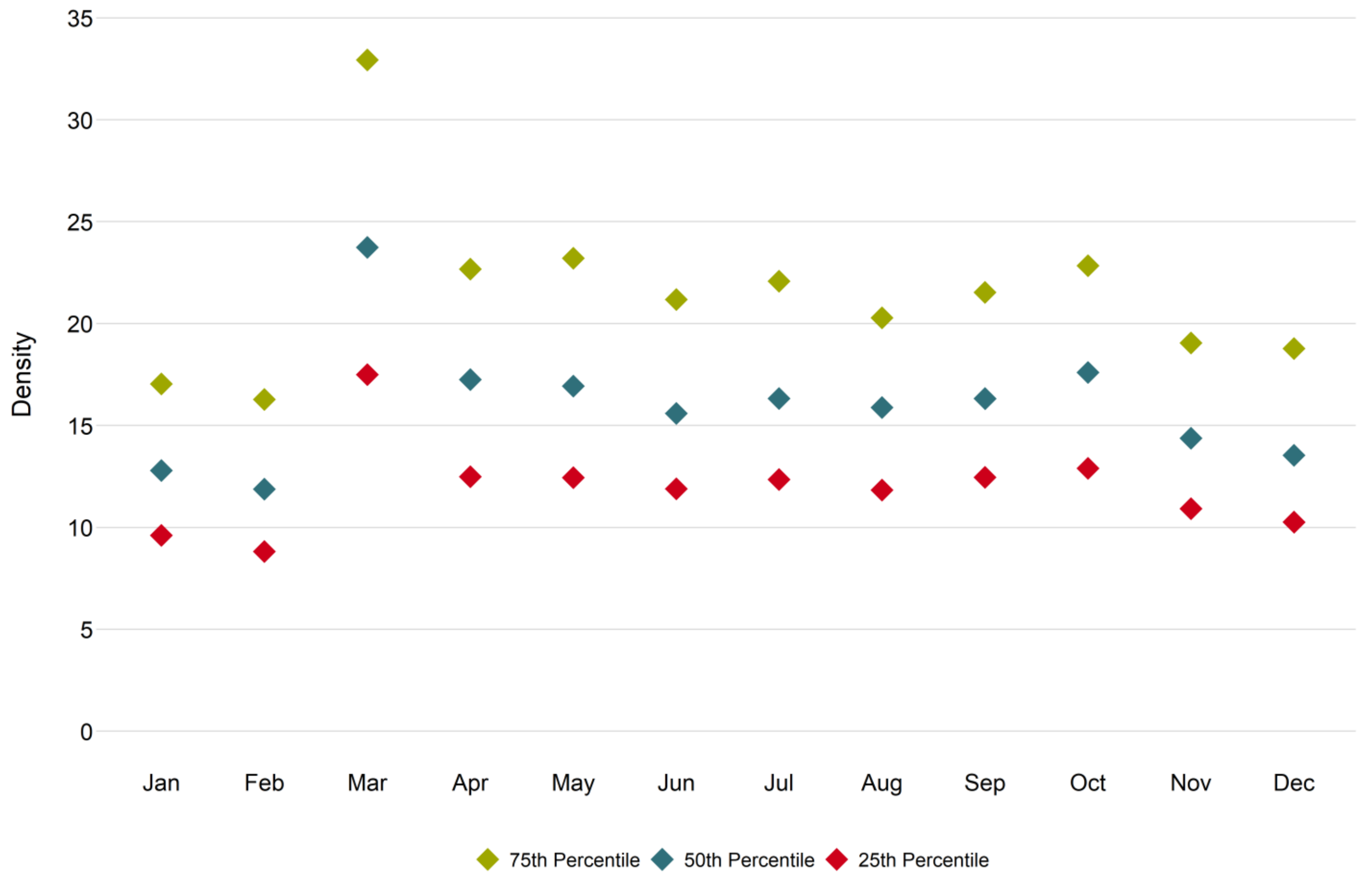


Source: Frontier Economics / UCL analysis using SERL data
Note: Black line indicates an approximation of the median; -0.0892%.

Error and bias by month

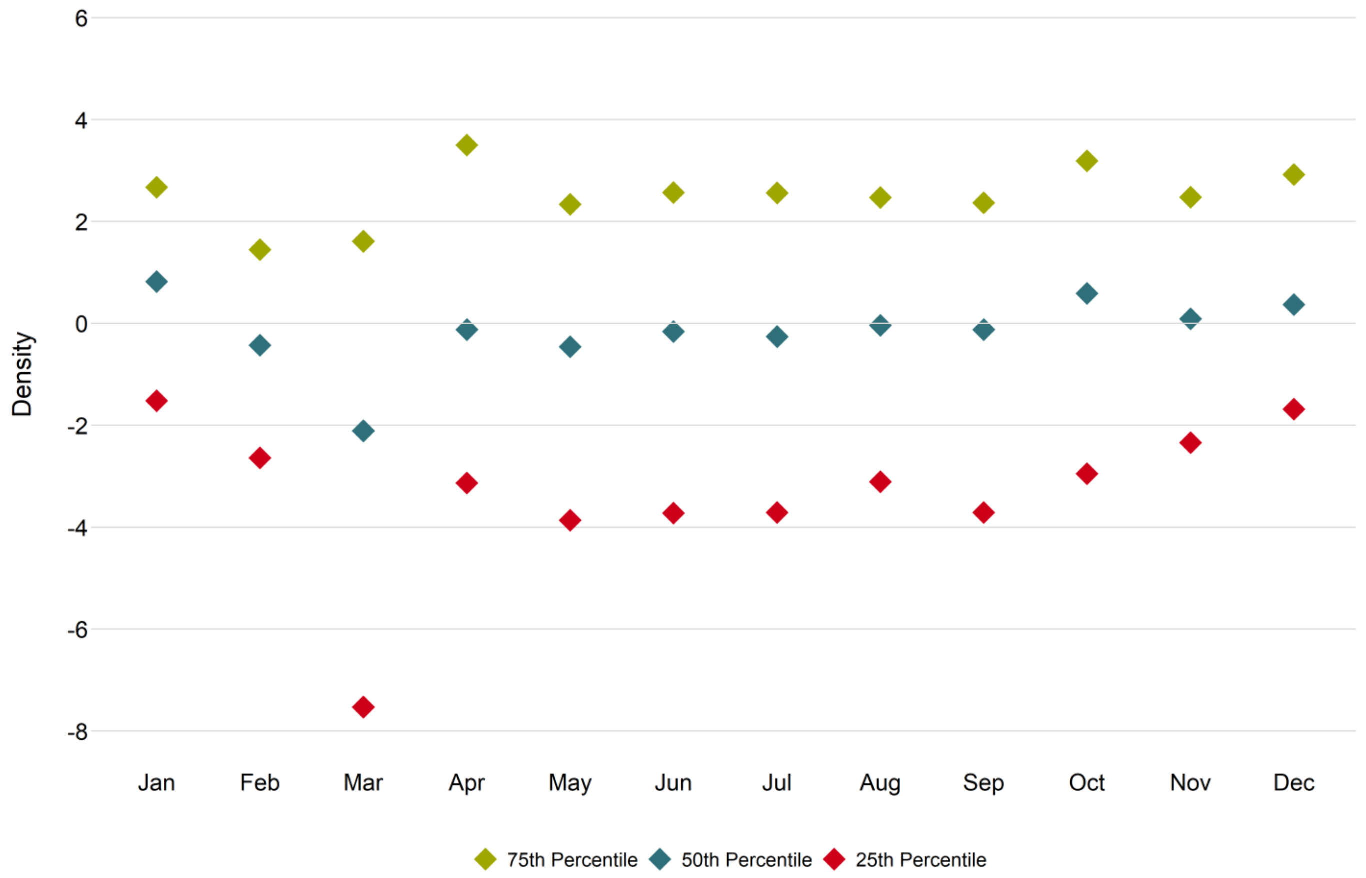
We can also consider the error and bias by month to see if the model has any seasonal weaknesses. Figure 39 shows similar error across all months except March. This is the month that was not required to be in the dataset in order not to reduce the sample size too much due to lack of data. So those with data in March may not have had much to make the predictions, and the errors are naturally larger. This is also evident in the subsequent figure which shows low bias for all months except March where there is a wider range of bias that is typically negative. Therefore care should be taken when interpreting results for March in future years, but otherwise we are not concerned with monthly error or bias.

Figure 39 Model error by month



Source: *Frontier Economics / UCL analysis using SERL data.*
Note: *Approximations for the median and interquartile range shown.*

Figure 40 Model bias by month



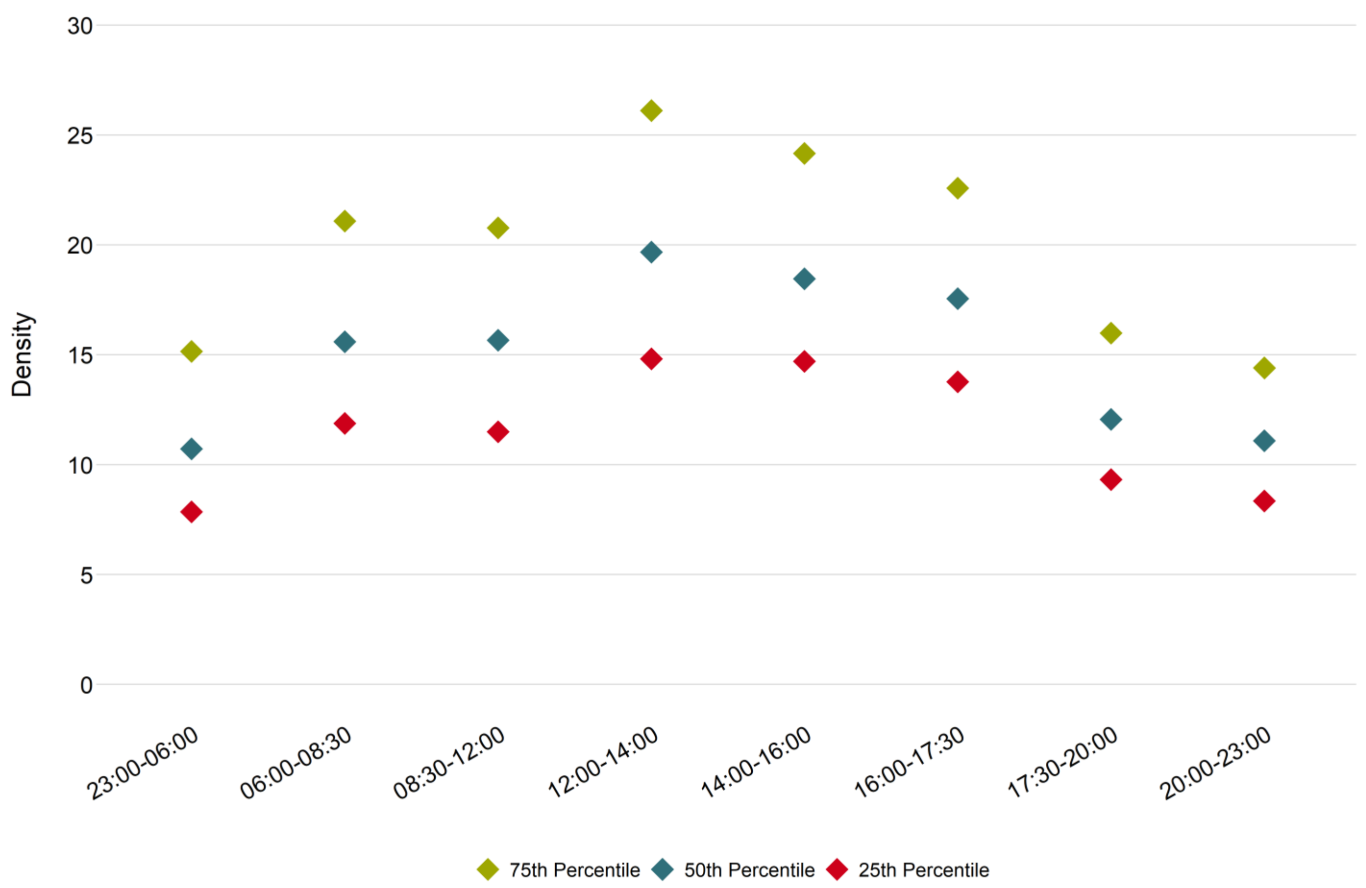
Source: Frontier Economics / UCL analysis using SERL data

Note: Approximations for the median and interquartile range shown

Error and bias by period of the day

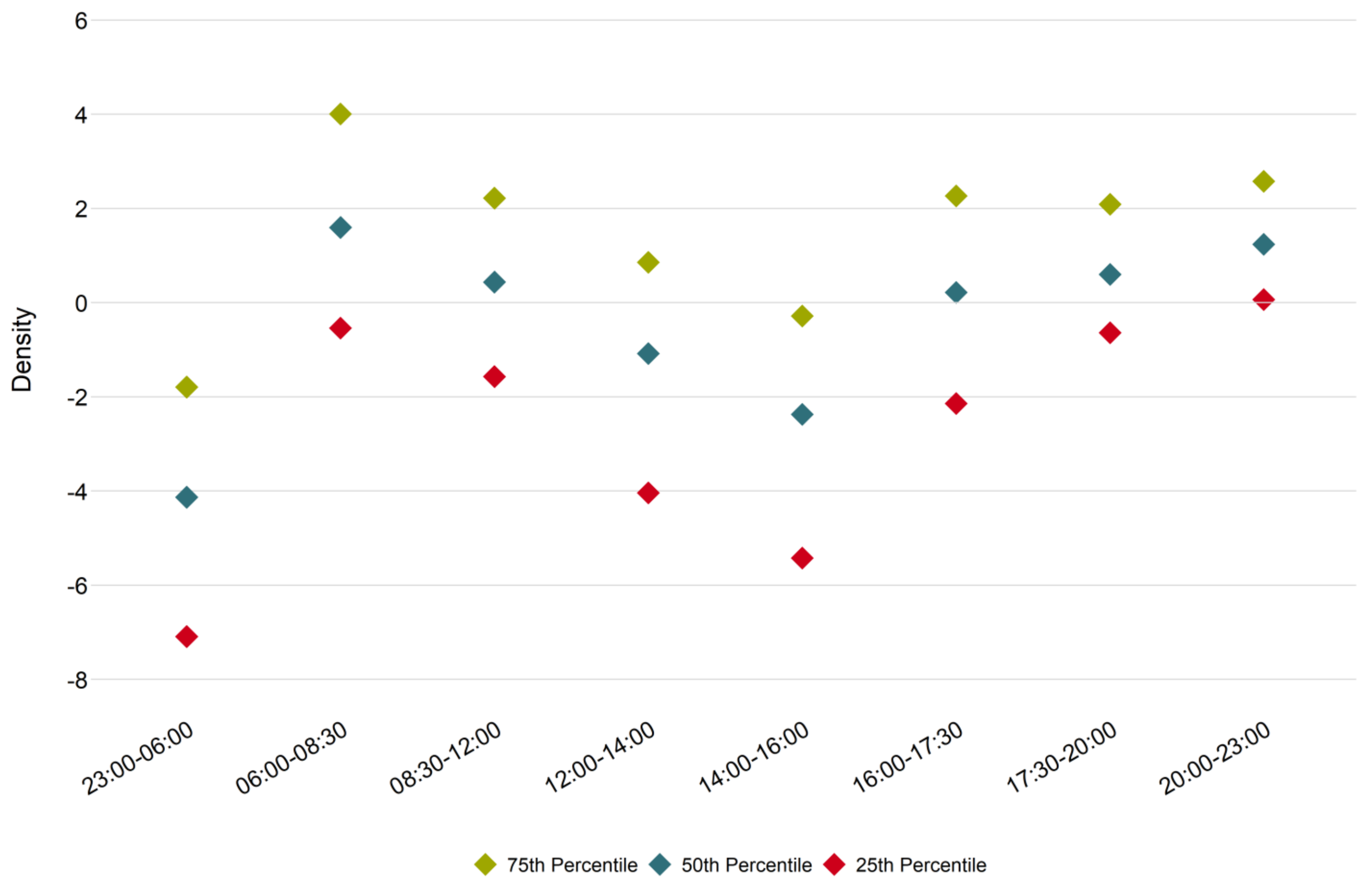
We see larger errors during the daytime than at night or during the evening, but across all periods errors remain fairly low. Bias for night-time demand prediction is negatively skewed, which is why modelling period A separately was considered, and for some households proved more accurate. Negative bias indicates under-prediction. It is likely that the models tend to find it harder to predict demand at the extreme ends of the scale, in particular at night when demand is typically very low. Care should be taken when interpreting results concerning period A, but as our analysis tends to focus on overall demand and peak demand, the bias we see here is not of particular concern.

Figure 41 Model error by period of the day



Source: Frontier Economics / UCL analysis using SERL data
Note: Approximations for the median and interquartile range shown

Figure 42 Model bias by period of the day



Source: *Frontier Economics / UCL analysis using SERL data*
Note: *Approximations for the median and interquartile range shown*

Considerations when using and interpreting model outputs and results

We have endeavoured to make the model predictions as accurate as possible by testing several models, modelling each house individually, and requiring 11 months of historic data (each month with at least 50% data availability) for training and testing. Cross-validation allowed us to perform model analysis of error and bias. We found that for most households model bias was very low, and model error was fairly low. However, there are a number of weaknesses that should be considered when working with the models and interpreting any results.

- Naturally some households were more predictable than others, likely due to household behavioural patterns and amount of data available. Some households suffered from much higher model error than others, and so it is best not to consider any household in isolation when comparing predictions with observations. The larger the group of households considered, the more confident we can be in our predictions.
- Models were trained on historic data up to the end of February 2020. This means that as we look further into the future our confidence in the model decreases, because households are more likely to have experienced significant changes since the training data period. For example, the arrival of

babies, children growing up and moving out, deaths within the household, change of working status, new pets, etc.

- The sample is small (586 households) and so it is difficult to generalise to the full GB population. We are aware of some of the ways in which it is and is not representative from our location and survey data, which should be considered when interpreting results. However it is not possible to capture all ways in which the sample is not representative of the broader population.

Despite the weaknesses mentioned above, the longitudinal data supported by detailed survey data has allowed us to create valuable predictive models that offer greater predictive power than is commonly done in other studies, and offer insights into how the COVID-19 pandemic has affected the electricity consumption of households in England and Wales.

Annex 3 – changes in consumption

This annex provides more details on the changes in consumption for different groups of customers, and how the change in consumption varies with changes in the number of COVID-19 cases.

First, we provide a more detailed set of results to show how the median change in consumption varies by household characteristics. This builds on the analysis in section 3.3.1. Next, we consider the relationship between the change in consumption and the COVID-19 incidence rate, by looking at the relationship between cases and the change in daytime consumption, and cases and the change in daily consumption. This builds on the analysis in section 4.1, where we look at the relationship between cases and the change in evening consumption.

Changes in consumption by customer characteristic

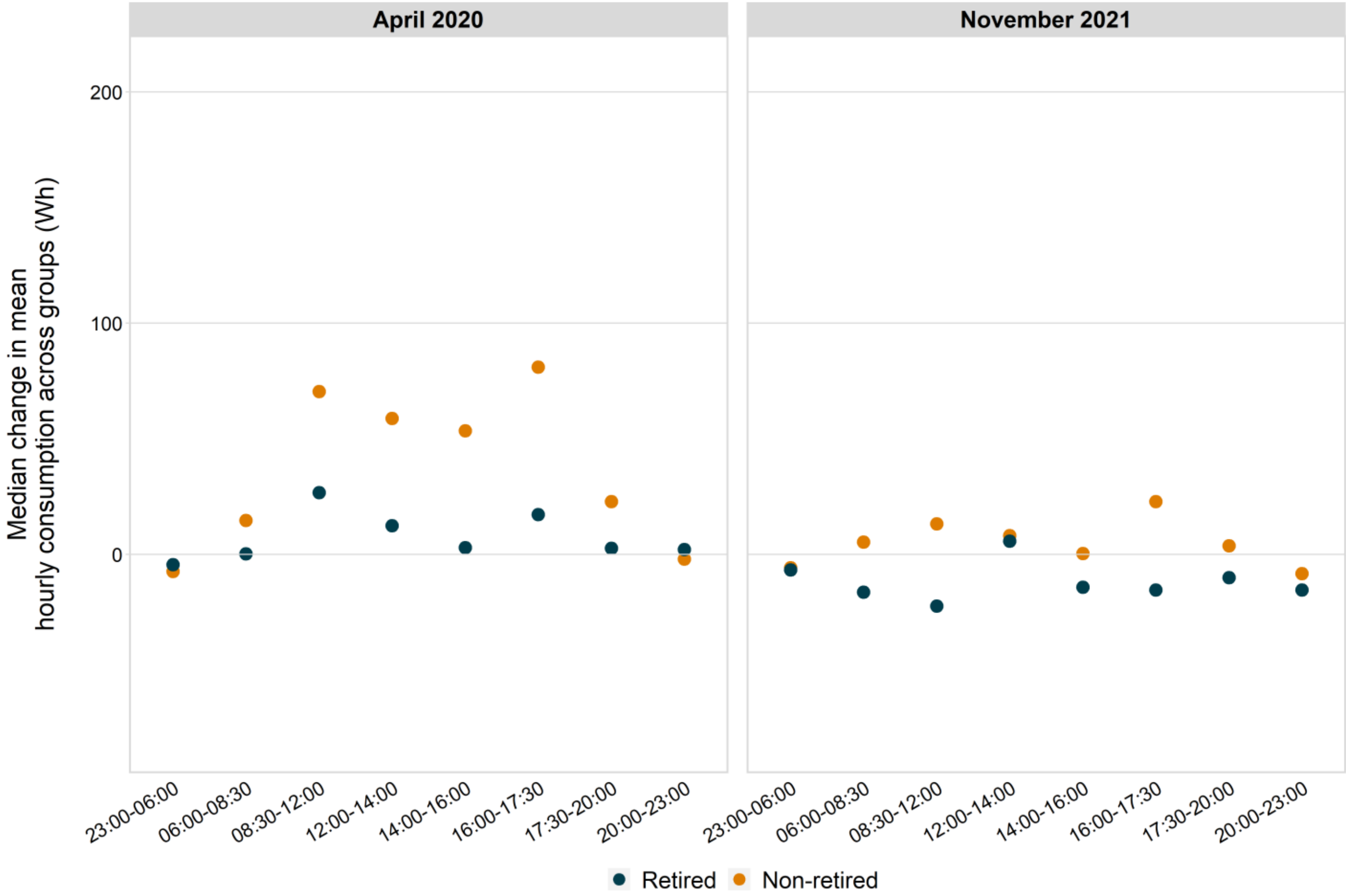
We assess the change in the approximate median household's load profile at the start and end of the sample, namely April 2020 and November 2021. We also use approximations for medians throughout this analysis by taking the average of at least 10 households with a value very close to the median, so that the value for an individual household is not revealed. It is worth noting throughout this analysis that there is significant variation across different characteristics, therefore we cannot generalise conclusions to all households with a specific characteristic. However it illustrates the broad differences between typical households. We consider the same household characteristics as assessed in 3.3.1.

Age of household

Retired households

The first group of customers we consider are retired households. Figure 43 shows how daytime and evening consumption change over time for households made up of retired members compared to other households.

Figure 43 Change in daily load profile in April 2020 and November 2021 – split by households with and without retired members



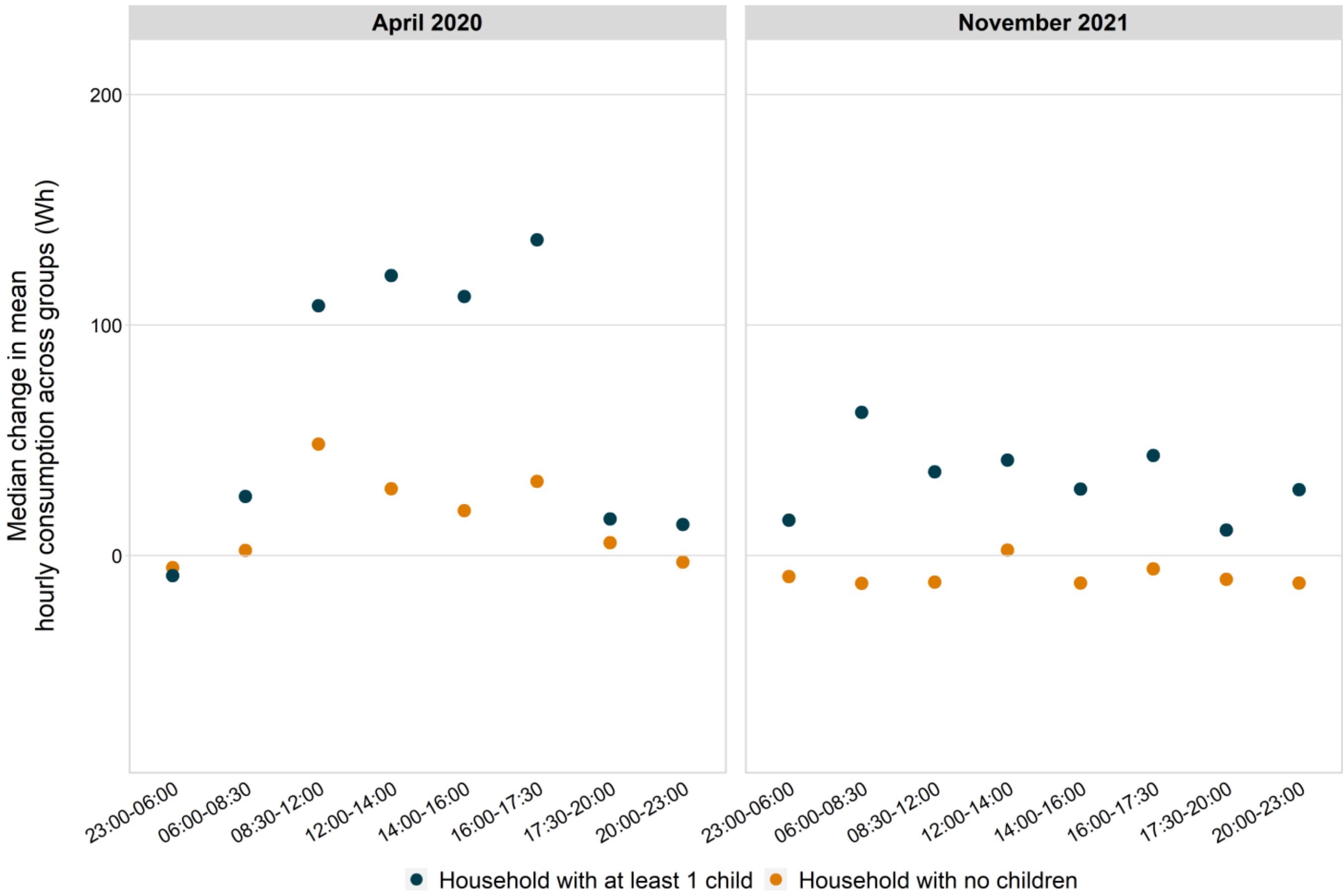
Source: Frontier Economics / UCL analysis using SERL data

The median household with retired members generally has small changes in their load profile in both April 2020 and November 2021. It has no change in consumption from the evening to early morning in April 2020, although they small increases in consumption in the morning and late afternoon in the same month. In November 2021, there is a small decrease in consumption across most of the day.

Households with children

We also assess the change in the daily load profile for households with and without children in Figure 44.

Figure 44 Change in daily load profile in April 2020 and November 2021 – split by households with and without children



Source: Frontier Economics / UCL analysis using SERL data

Households with at least one child have large increases in consumption across the middle of the day in April 2020, whereas they have little-to-no change in consumption in the evening and night during the same month. Households without children have a similar pattern of differences in April 2020, however the change in load profile is flatter, with smaller changes in consumption across the middle of the day.

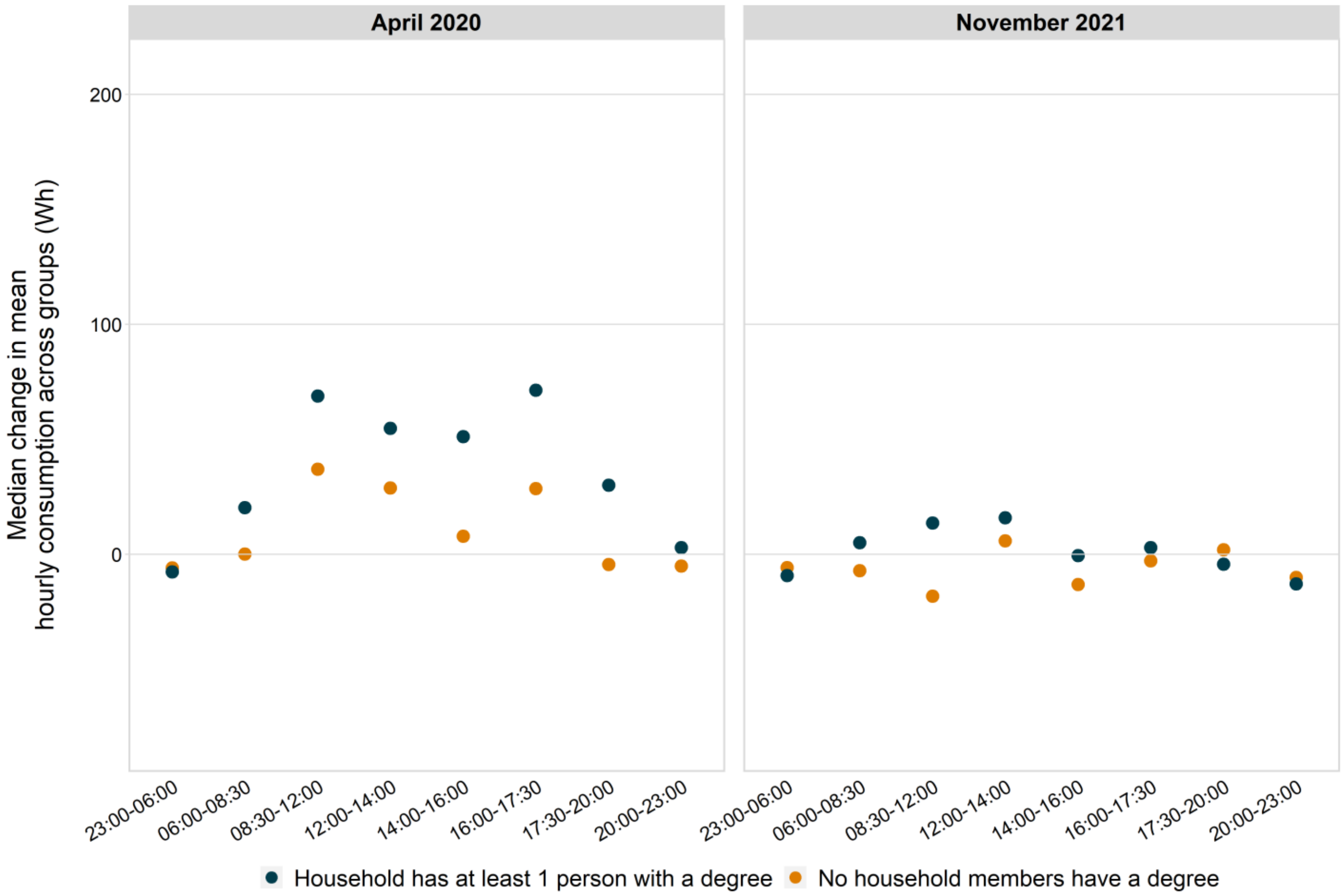
The change in consumption for households with children in November 2021 is notably different to April 2020, with smaller increases across the middle of the day, but slightly larger increases across the night, and a considerably larger increase from 06:00 – 08:30. Households without children have a small decrease in consumption across most of the day in November 2021, which is also different to the pattern we see for the same type of household in April 2020.

Employment status and homeworking

Households where at least one occupant has a degree

In Figure 45 we consider the change in the daily load profile for households where at least one occupant has a degree.

Figure 45 Change in daily load profile in April 2020 and November 2021 – split by households with and without at least one member who has a degree



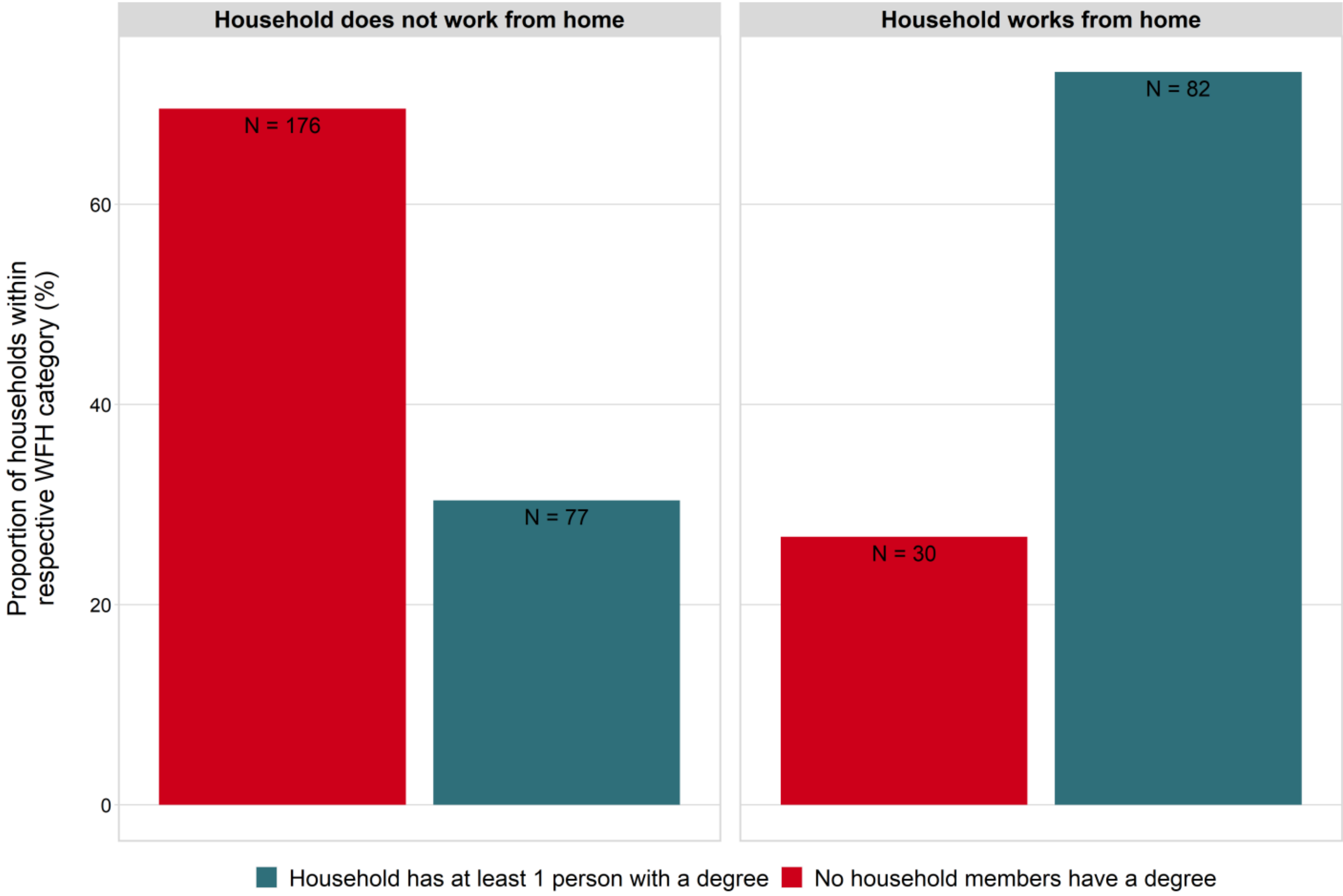
Source: Frontier Economics / UCL analysis using SERL data

Both types of households have a similar change in daily load profile in April 2020, except that the changes in consumption are larger for households where at least one occupant has a degree compared to households where no occupants have a degree. Exceptions are the sub-daily periods 06:00 – 08:30 and 17:30 – 20:00, when households where no occupants have a degree have no change in consumption, whereas households where at least one occupant has a degree have an increase in consumption.

Both types of households have a different pattern in November 2021. Households where no occupants have a degree have no-to-little changes in consumption, with small decreases from the night to the morning and mid-afternoon, and a small increase in the early afternoon. Households where at least one occupant has a degree have small increases form the morning until the early afternoon, little change until the evening, and then a small decrease across the night.

In Figure 46 we consider how the likelihood of a household having at least one occupant with a degree varies depending on whether the household works from home.

Figure 46 Likelihood of a household having a degree holding occupant for different working from home statuses



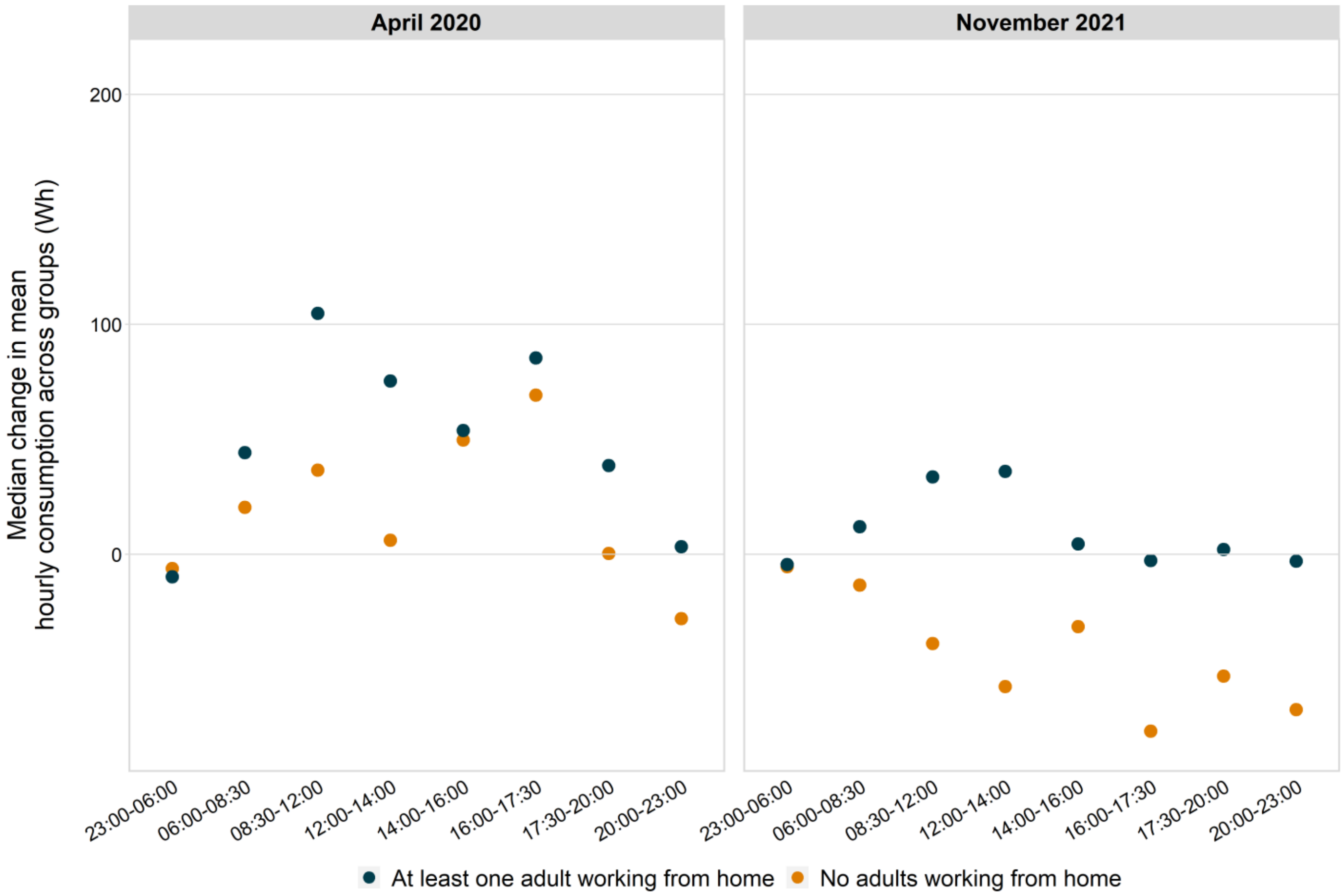
Source: Frontier Economics analysis

The above chart shows that, for households that don't work from home, 70% don't have at least one occupant with a degree. This means that 30% of households don't work from home and have at least one occupant with a degree. For households that do work from home, 73% have at least one occupant with a degree, and so 27% are households that work from home but don't have a degree.

COVID-19 survey on homeworking

Figure 47 below shows the change in consumption for households where at least one member worked from home. See section 3.3.1 for details of how we have defined homeworking.

Figure 47 Change in daily load profile in April 2020 and November 2021 – split by households with and without at least one member who worked from home



Source: Frontier Economics / UCL analysis using SERL data

In April 2020, households with at least one member working from home have large increases in consumption across most of the day, from 06:00 – 20:00, and they have little-to-no change in consumption in the evening and night during the same month. Households with at no members working from home have increases from 06:00 – 12:00 and 14:00 – 17:30, however these increases are smaller than the increase for households with at least one member working from home.

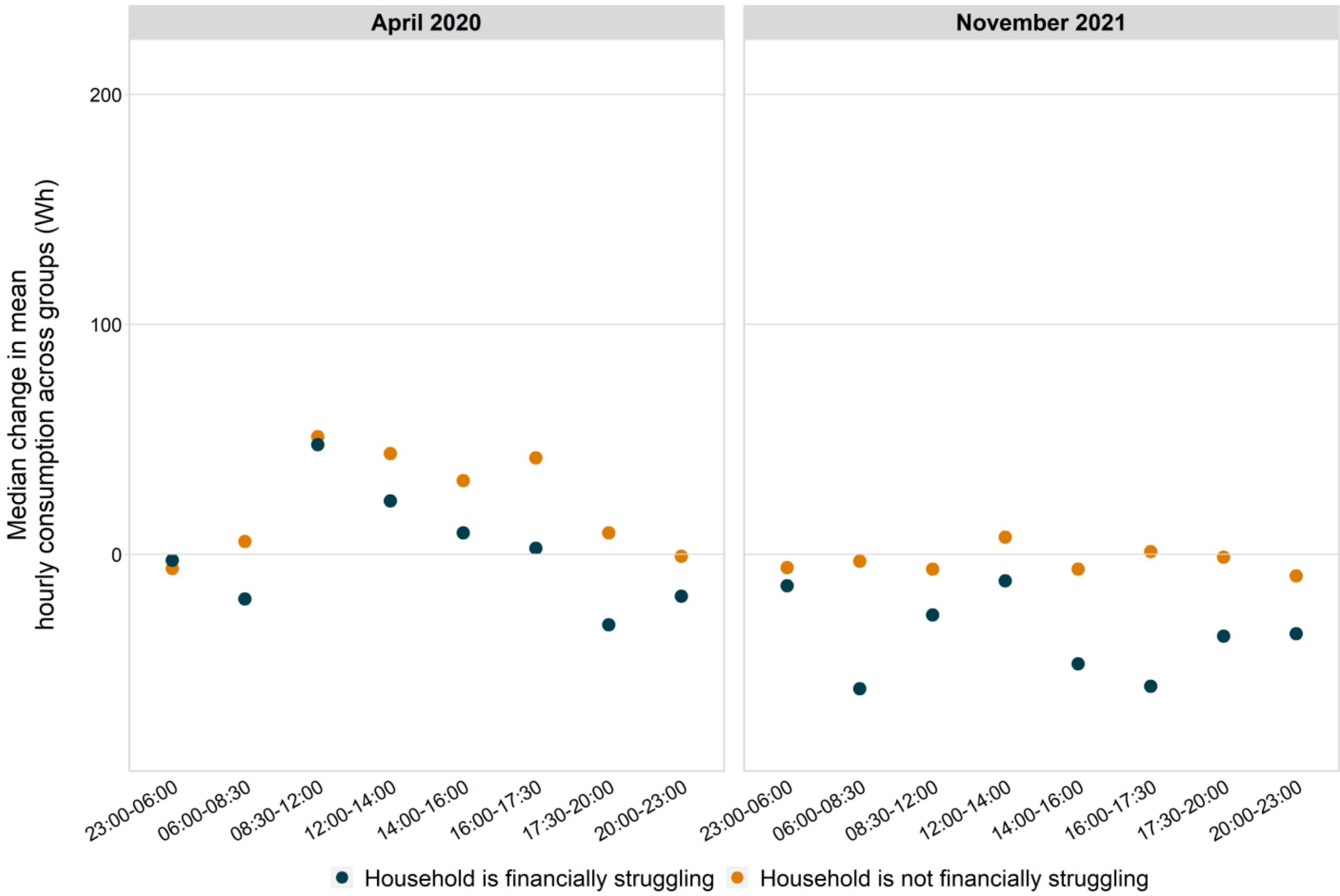
In November 2021, both types of households have less positive changes in consumption compared to April 2020. Households with at least one member working from home have small-to-moderate increases in consumption between 06:00 – 14:00, but outside of this period changes in consumption are minimal. Households with at no members working from home have moderate decreases in consumption from 06:00 – 20:00, with particularly large decreases from 12:00 – 14:00, and 17:30 – 23:00. The large decreases could be due to November 2021 being almost two years from the training data, which can decrease our confidence in the model as households are more likely to have experienced significant changes since the training data period (e.g. births, deaths, people moving out). The decreases may also be explained to some extent by other unobservable factors such as changes in behaviour.

Income

How well are survey respondents managing financially

Figure 48 below shows the change in consumption for household split by whether they self-reported as 'managing financially'.

Figure 48 Change in daily load profile in April 2020 and November 2021 – split by households which are and aren't 'managing financially'



Source: Frontier Economics / UCL analysis using SERL data

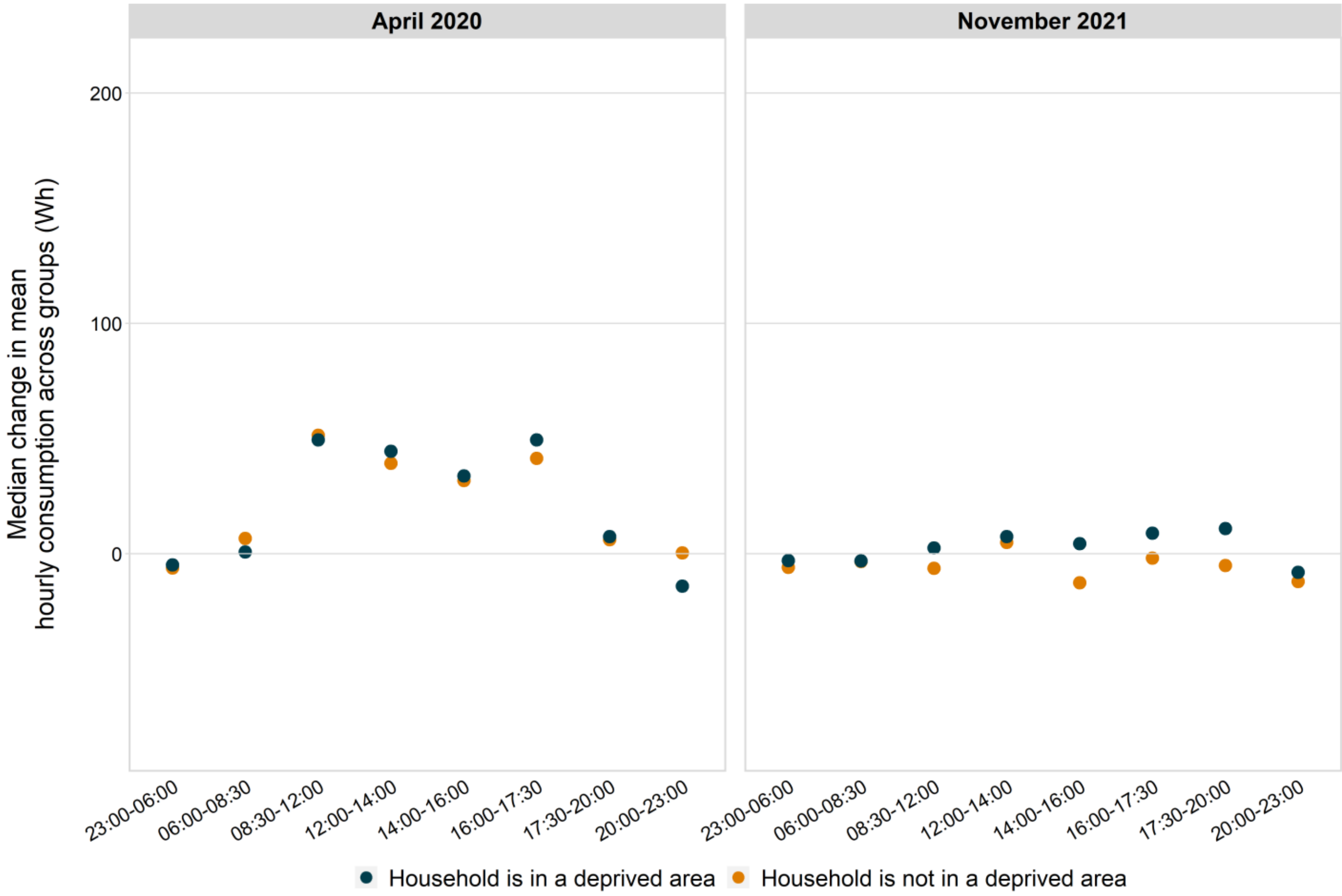
In April 2020, households which reported to be financially struggling typically have a lower change in load profile than household which did not. The largest increase for financially struggling households occurs during 08:30 – 12:00, and they also have small increases from 12:00 – 16:00. The largest increase for households not financially struggling is also during 08:30 – 12:00, but these households have relatively larger increases in consumption from 12:00 – 17:30 than households which are struggling.

In November 2021, households which aren't financially struggling have little change in consumption across the day. On the other hand, financially struggling households have decreases in consumption across most periods of the day, with the largest decreases occurring during 06:00 – 08:30, and 14:00 – 17:30. As mentioned above, the large decreases could be due to unobservable factors such as the household size changing, and changes in behaviour.

Index of multiple deprivation

Figure 49 below shows the change in consumption for household split by whether the household is located in an area of deprivation.

Figure 49 Change in daily load profile in April 2020 and November 2021 – split by households which are and aren't located in a deprived area



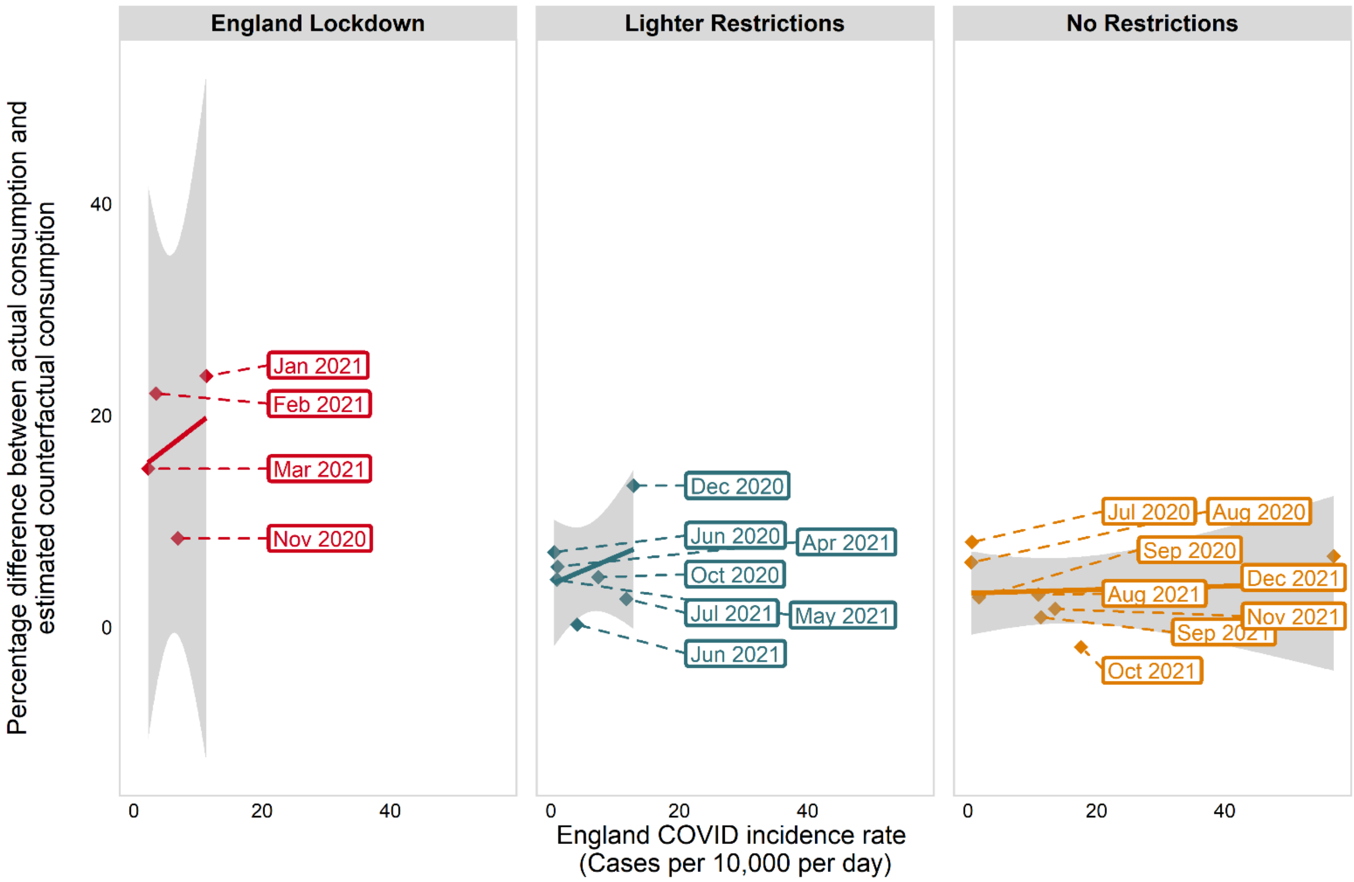
Frontier Economics / UCL analysis using SERL data

Across both months, both types of households follow a similar pattern. Both types of households have increases in consumption from 08:30 – 17:30 during April 2020. In addition, both types of households have a relatively flat change in load profile during November 2021, with households in a deprived area having a slighter greater increase in consumption then households not located in a deprived area.

Changes in consumption and COVID-19 cases

This subsection considers the relationship between the change in consumption and the COVID-19 incidence rate. Figure 26 in the main report shows the relationship between the change in evening consumption and estimated COVID-19 cases in England. Below we show the relationship for daytime and total consumption.

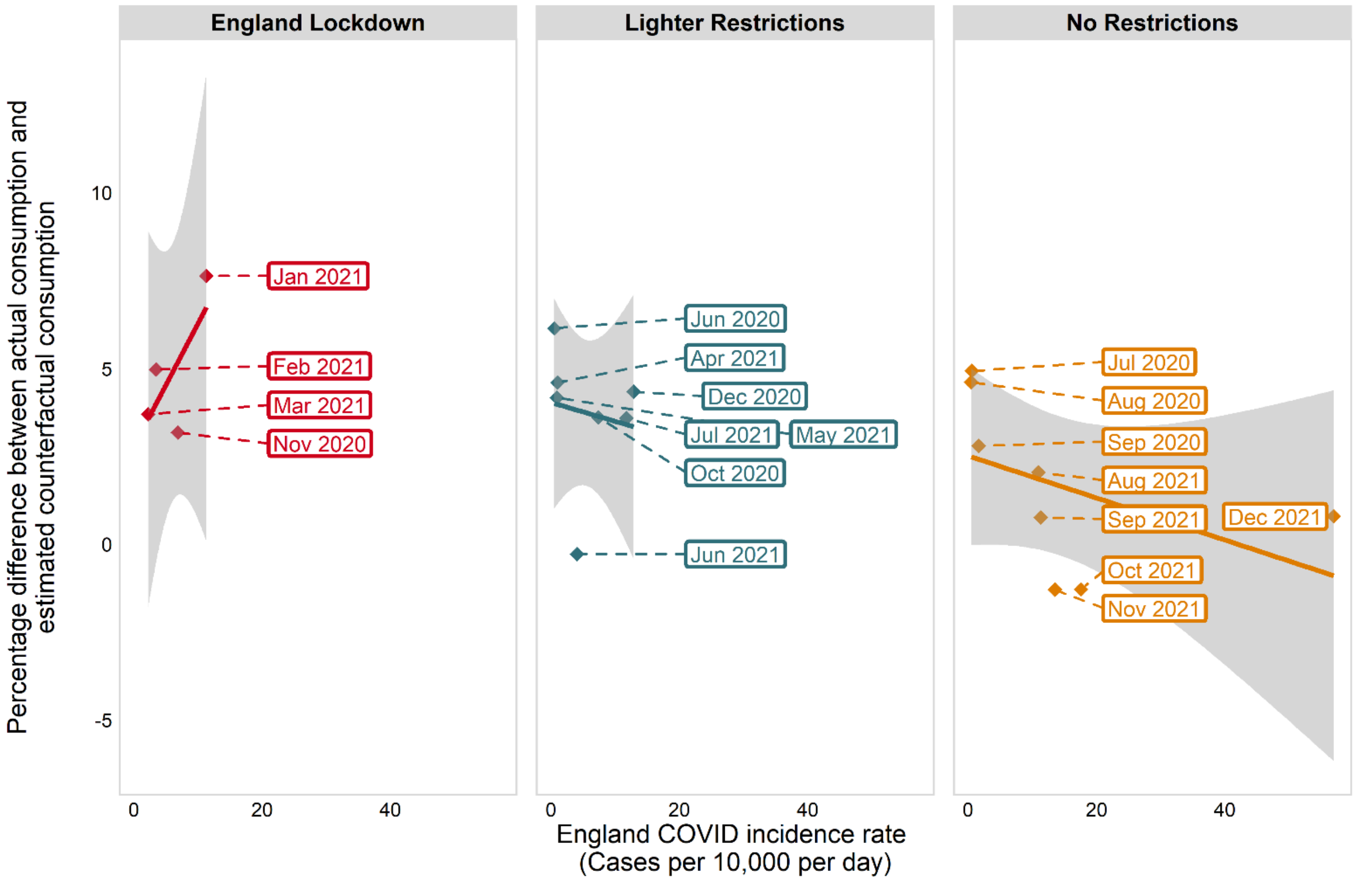
Figure 50 Relationship between estimated cases in England and changes in daytime consumption



Source: *Frontier Economics analysis of SERL data*

Figure 50 above shows that there is a positive correlation between estimated cases and the change in daytime consumption for all three types of restriction. As expected, lockdown months typically have higher levels of changes in consumption compared to lighter and no restrictions, as this is likely associated with greater behaviour change. However, within each category of months, the relationship between case numbers and consumption change is very weak.

Figure 51 Relationship between estimated cases in England and changes in consumption for all periods of the day



Source: Frontier Economics analysis of SERL data

Figure 51 above shows that there is a positive correlation between estimated cases and the change in consumption for months of full lockdown. On the other hand, it shows that there is a negative correlation between estimated cases and the change in consumption for months where there were either lighter or no restrictions. This appears to be driven by later periods having higher numbers of COVID-19 cases but lower changes in consumption. This might be driven by the impact of vaccinations and increased "lockdown fatigue".

Regression analysis

Overview of outputs

The analysis focuses on how different customers were affected by the pandemic, and how that changed by time of the day, and how that changed over time. We consider this by running regressions to analyse the relationship between changes in consumption as a result of the pandemic and household demographics. We use regression analysis to understand the relative impact of multiple variables on a given measure. In particular, we want to find out whether there are relationships between household characteristics and consumption changes, whether the relationships are independent and whether there is statistical significance associated with them.

In the main report, we assess the changes in electricity consumption associated with each individual characteristic separately. However, households will have multiple characteristics which may be associated with different changes in electricity consumption. There will likely be some household characteristics that are associated with other household characteristics. Using regression analysis, we can start to unpick which characteristic is associated with changes in electricity consumption. The correlation matrix below in Table 7 illustrates this.

We carry out regressions in four time periods:

- April 2020, period D (12:00-14:00);
- April 2020, period G (17:30-20:00);
- November 2021, period D (12:00-14:00); and
- November 2021, period G (17:30-20:00).

For each of the regressions, we have the same dependent variable. This is the calculated change in consumption between the modelled counterfactual and observed consumption. However, we normalize this change in consumption by calculating the consumption as a proportion of annual consumption for that household. We do this in order to normalize for the overall size of household consumption in a particular year. This analysis then loses information about the overall size of consumption changes, but helps compare types of changes seen by customers. Given that we are interested in changes in consumption, we adjust average consumption levels for each household to the Typical Domestic Consumption Value (TDCV)[79]. This is the same methodology as for the clustering analysis in Section 3.3.2.

For each of the regressions, we have the same independent variables across the regressions, these are:

- An indicator variable (i.e. one that is always zero or one) for households made up of retirees;
- an indicator variable for households with at least one child;
- an indicator variable for households that responded to survey signalling that they were financially struggling in some way; and
- an indicator variable for households with at least one person with a degree.

Below we show the same outputs for regressions in each time period. The outputs are all formatted the same:

- The 'variable' relates to the independent variables listed above that we are testing against the dependent variable;

[79] The TDCV (as set by Ofgem) for electricity is 2,900 kWh for 2021. This is set out in the 'Typical Domestic Consumption 2021 Decision Letter': https://www.ofgem.gov.uk/sites/default/files/docs/2021/05/tdcv_decision_letter_2021_0.pdf

- the ‘estimate’ variable is the coefficient estimate from the regression;
- the ‘std error’ variable is the standard error of the coefficient estimate;
- the ‘t-stat’ variable is the result of the t-test for each coefficient estimate;
- the ‘p-value’ variable is the resulting p-values for each coefficient estimate; and
- the ‘significance’ variable shows whether each coefficient is significant and at what probability, more specifically:
  - One star ‘*’ indicates significance at the 10% level;
  - Two stars ‘**’ indicates significance at the 5% level;
  - Three stars ‘***’ indicates significance at the 1% level; and
  - No stars indicates no statistically significant difference from zero.

We note that this is a relatively unsophisticated analysis to illustrate whether different household characteristics are associated with different responses in electricity consumption during the pandemic. These are not intended to be comprehensive explainers of all changes in consumption, which is reflected in the low R-squared values.

Since there may be some level of multicollinearity between the independent variables, for each regression we take a correlation matrix of the independent variables. This will help us understand the precision of the estimated coefficients. If there is multicollinearity present, then that may reduce the statistical power of the regression model. Doing this will help us understand the extent to which this impacts our modelling. The stars indicate significance levels as per the list above.

Correlation matrix

Table 7 shows correlations between independent variables for April 2020, period D (the daytime period). This suggests the following effects:

- Whether or not the household is financially struggling is not significantly correlated to any other characteristics.
- If the household is retired, the correlation matrix implies there is some correlation with households without children and correlation with households without degrees. This suggests retired households tend not to have children or members with degrees.
- For households with children, there is a small positive correlation with households that have a degree. This suggests households with children are also more likely to have members with degrees.

Table 7 Correlation matrix results

Correlations	financially struggling dummy	household retired dummy	child dummy	degree dummy
financially struggling dummy	-			
household retired dummy	-0.05	-		
child dummy	0.01	-0.37***	-	
degree dummy	-0.02	-0.31***	0.17***	-

Source: Frontier Economics analysis of SERL data
Note: N=560

This shows that a simple analysis of the changes in consumption by household characteristic could be affected by multicollinearity. The following regression analysis can start to pick apart these effects and identify the underlying drivers of changes in consumption in different periods.

April 2020, period D

Table 8 shows the regression results for April 2020, Period D (the daytime period). The results suggest that if the household has a child or someone in the household has a degree, then they will have a positive and statistically significant (at the 1% and 10% level respectively) difference in the change in daytime consumption as a result of the pandemic versus those without a child or degree. Conversely, if the household is retired then they will have a negative and statistically significant (at the 5% level) difference in the change in daytime consumption as a result of the pandemic versus those who aren't retired. Whether or not a household is financially struggling appears to make no statistically significant difference to changes in consumption.

Table 8 Regression 1 results – April 2020, daytime

variable	estimate	std.error	t-stat	p.value	significance
(Intercept)	6.270921	1.490745	4.206568	3.02E-05	***
household retired dummy	-3.55878	1.707179	-2.0846	0.037562	**
child dummy	6.171982	2.26029	2.730615	0.006523	***
financially struggling dummy	-0.87049	3.814985	-0.22818	0.819592	
degree dummy	2.846049	1.600668	1.778038	0.075945	*

Source: Frontier Economics analysis of SERL data
Note: N=560, Multiple R-sq = 0.04691, Adj R-sq = 0.04004

April 2020, period G

Table 9 shows the regression results for April 2020, Period G (the evening peak period). The results suggest that if someone in the household has a degree, then they will have a positive and statistically significant (at the 10% level) difference in the change in evening consumption as a result of the pandemic versus those without a degree. All of the other variables measured are not statistically significant.

Table 9 Regression 2 results – April 2020, evening peak

variable	estimate	std.error	t-stat	p.value	significance
(Intercept)	2.44782	2.214584	1.105318	0.269498	
household retired dummy	-3.15753	2.541273	-1.2425	0.214574	
child dummy	-3.12305	3.352818	-0.93147	0.352013	
financially struggling dummy	-3.18869	5.690931	-0.56031	0.575493	
degree dummy	5.464693	2.381158	2.294973	0.022105	*

Source: Frontier Economics analysis of SERL data

Note: N=563, Multiple R-sq = 0.01642, Adj R-sq = 0.009365

November 2021, period D

Table 10 shows the regression results for November 2021, Period D (the daytime period). None of the coefficients are significant, suggesting that none of the independent variables make a significant difference in explaining daytime consumption changes across households during the pandemic. This is unsurprising, given that the impact of the pandemic is likely to have faded over time.

Table 10 Regression 3 results – November 2021, daytime

variable	estimate	std.error	t-stat	p.value	significance
(Intercept)	1.057908	1.344545	0.786815	1.057908	
household retired dummy	-0.82333	1.552	-0.5305	-0.82333	
child dummy	0.802754	2.042949	0.392939	0.802754	
financially struggling dummy	-2.77973	3.416726	-0.81357	-2.77973	
degree dummy	1.794648	1.453372	1.234817	1.794648	

Source: Frontier Economics analysis of SERL data

Note: N=516, Multiple R-sq = 0.007214, Adj R-sq = - 0.0005575

November 2021, period G

Table 11 shows the regression results for November 2021, Period G (the evening peak period). None of the coefficients are significant, suggesting that none of the independent variables make a significant difference in explaining evening consumption changes across households during the pandemic. This is unsurprising, given that the impact of the pandemic is likely to have faded over time.

Table 11 Regression 4 results – November 2021, evening peak

variable	estimate	std.error	t-stat	p.value	significance
(Intercept)	0.894392	1.885564	0.474336	0.635462	
household retired dummy	-0.33949	2.169207	-0.1565	0.875697	
child dummy	0.833164	2.847595	0.292585	0.769958	
financially struggling dummy	1.732667	4.776858	0.362721	0.716963	
degree dummy	-1.60621	2.029975	-0.79125	0.429166	

Source: Frontier Economics analysis of SERL data

Note: N=517, Multiple R-sq = 0.001656, Adj R-sq = -0.006143

Conclusions

We can draw some broad conclusions from the regression analysis to support our conclusions in Section 3.

- Households without retired members had some positive changes during the pandemic. During the initial lockdown period, this appeared to be associated with the presence of retirees, particularly for changes in daytime consumption. However, the difference in changes in evening consumption during the initial lockdown were more likely to be driven by other factors, suggesting multicollinearity between these households and households with a degree. The regression analysis suggests differences in changes in consumption between households with and without retired members were not due to that household characteristic.

- Households with children also had some positive changes during the pandemic, exhibiting similarities to households without retired members (which is perhaps unsurprising, given the strong correlation between the variables). Households with children was associated with changes in daytime electricity consumption during the initial lockdown, but not with changes in evening consumption during the same period. Changes in evening consumption was more likely to have been driven by a correlated characteristic: households with a degree. This characteristic was not associated with changes in consumption during any period in November 2021.

- There is no evidence of any statistical relationship between changes in electricity consumption and those struggling financially. This aligns with our conclusions.

- There is some evidence of households with a degree being associated with greater changes in consumption during the pandemic compared with other households. This is true for both daytime and evening consumption. However, there is no statistically significant effect in November 2021.

Annex 4 – Clustering analysis

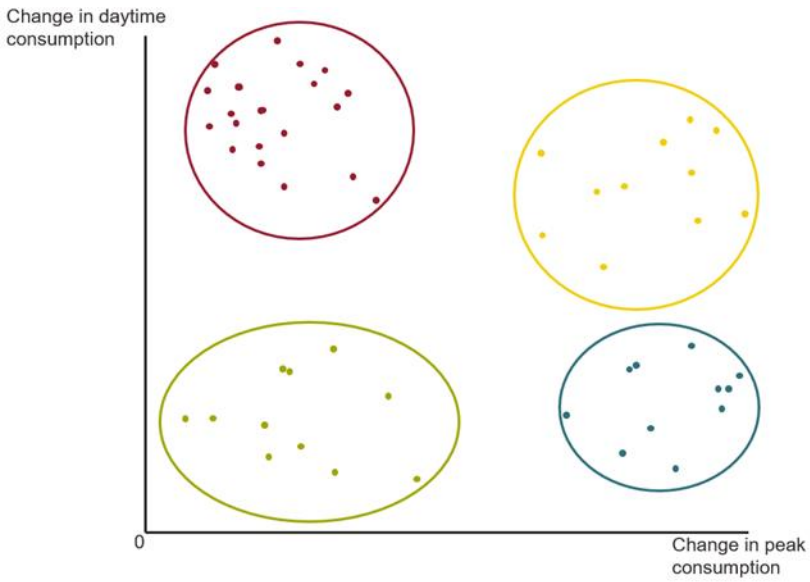
Section 3.3.2 presents the results of a clustering analysis which groups households based on the changes in their consumption across six months. In this annex we provide further information on:

- the clustering algorithm we use, the method we use to select the optimal number of clusters, and the data cleaning and preparation before the clustering analysis is undertaken;
- additional sensitivities (setting the number of clusters to two, and pooling results from across adjacent months); and
- results for a simpler version of the analysis considering only a single month at a time (which allows the results to be shown more graphically).

Clustering methodology and data cleaning

Clustering analysis allows us to group customers based on how similar they are across a set of observations. A simple, illustrative example is shown in Figure 52. Each point on the chart represents an individual household: the x-axis showing the change in evening peak consumption and the y-axis showing the change in daytime consumption. The clustering algorithm groups together these points into a pre-determined number of clusters based on how close the observations are together. In this example, we determined four clusters, and the circles drawn round each set of points indicates which observations belong to which cluster.

Figure 52 Illustrative changes in consumption for two sub-daily periods in a given month



Source: Frontier Economics – illustrative clustering analysis

The clustering algorithm

The clustering algorithm we use to group observations into clusters is called **K-means clustering**. First, the algorithm randomly picks centroids (ie centres of clusters) for ‘K’ number of clusters. All observations are then assigned to their nearest cluster, and the centroid is moved to the average of the observations assigned to that centroid. As the centroid moves, the algorithm re-assigns observations to the closest centroid. The algorithm repeats the previous steps until the sum of distances between observations and their centroid is minimised. The value of the centroid is then the typical consumption of households within that cluster.

Data cleaning

Before running the K-means clustering algorithm, we clean and prepare the data by:

- **Removing erroneous observations**. We remove any observations where the estimated counterfactual consumption is negative.
- **Normalising the change in consumption variable**. When considering the average change in consumption, we do not want the analysis to be impacted by relatively small customers who might have a relatively large percentage change in demand, but only a small change in actual consumption (or vice versa for relatively large customers). Hence, for each household, we adjust each change in consumption observation by expressing each observation as a proportion of annual estimated counterfactual consumption.
- **Removing outliers**. We remove any observations where the normalised change in consumption, which is also expressed as a proportion of annual estimated counterfactual consumption, is greater than 3% or less than -3%. This is based on outlier analysis which suggests that ±3% are reasonable thresholds to exclude observations.
- **Scaling each observation to the level of the typical consumer**. We multiply each observation by 2900, which is Ofgem's 2021 value for the annual medium Typical Domestic Consumption Value (TDCV) for electricity profile class 1.[80]
- **Removing households that have missing observations**. K-Means clustering requires a dataset with no missing values, and so any households with one or more missing consumption observations will be dropped from the dataset.

Determining the optimal number of clusters

To determine the optimal number of clusters for the clustering analyses, we calculate the silhouette coefficient for each cluster, which is a measure of the similarity of an observation to its own cluster compared to other clusters and can range from -1 to 1. A negative silhouette coefficient suggests that an observation should belong to another cluster rather than its own cluster, whilst a positive silhouette coefficient suggests that an observation should remain in its currently assigned cluster. The more positive the silhouette coefficient, the better the assignment of observations to clusters given the number of clusters.

The overall silhouette score is the mean of the silhouette coefficient, and the number of clusters corresponding to the highest silhouette score is often chosen as the number of clusters in our analyses.

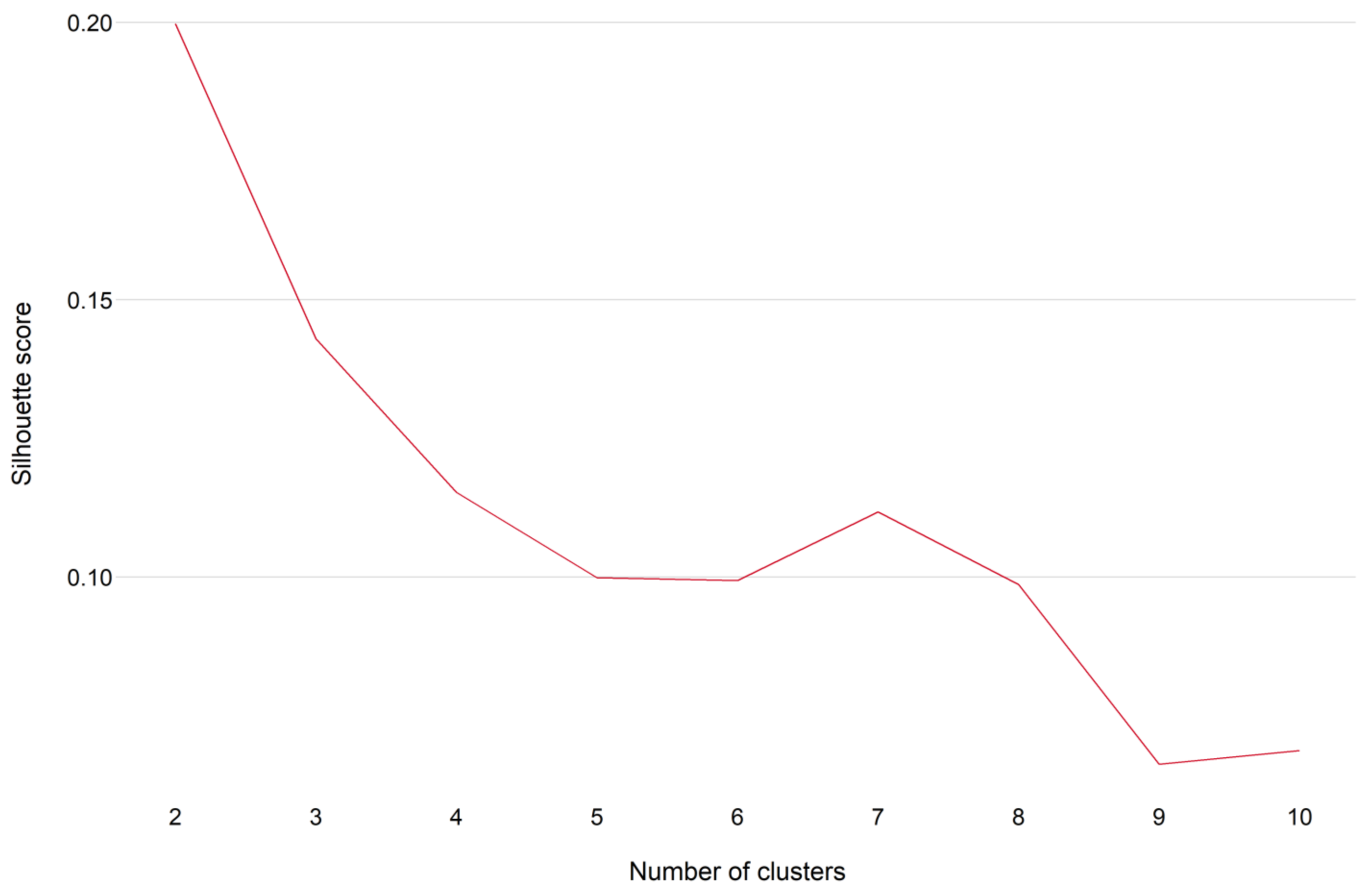
[80] Ofgem 2021, Typical Domestic Consumption Values 2021 Decision Letter. Available at: https://www.ofgem.gov.uk/publications/decision-typical-domestic-consumption-values-2021

Two cluster analysis

Section 3.3.2 presents the results for a clustering exercise where seven clusters have been specified. The silhouette score for seven clusters is a 'local maximum' (i.e. higher than the score for six or eight clusters), however the highest score was for two clusters. Despite this, seven clusters was used in order to provide a more fine-grained set of results which can help distinguish different patterns of consumption. Here, we provide the results of the version of this analysis using only two clusters.

Figure 53 below shows the silhouette score, including the 'local maximum' described above.

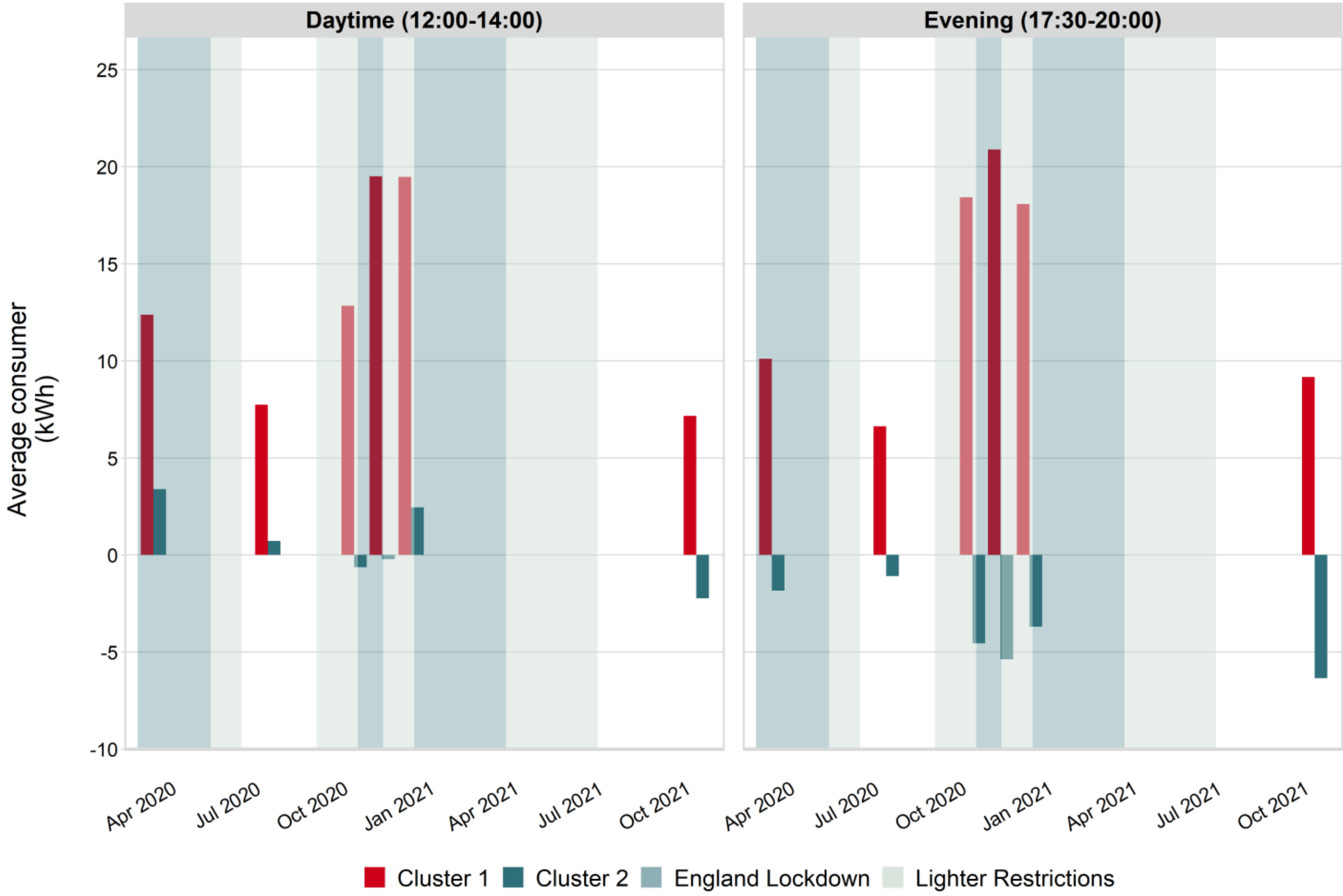
Figure 53 Silhouette score for non-pooled analysis



Source: Frontier Economics / UCL analysis using SERL data

Error! Reference source not found.

Figure 54 Change in daytime and evening consumption over time by cluster for 2 cluster analysis

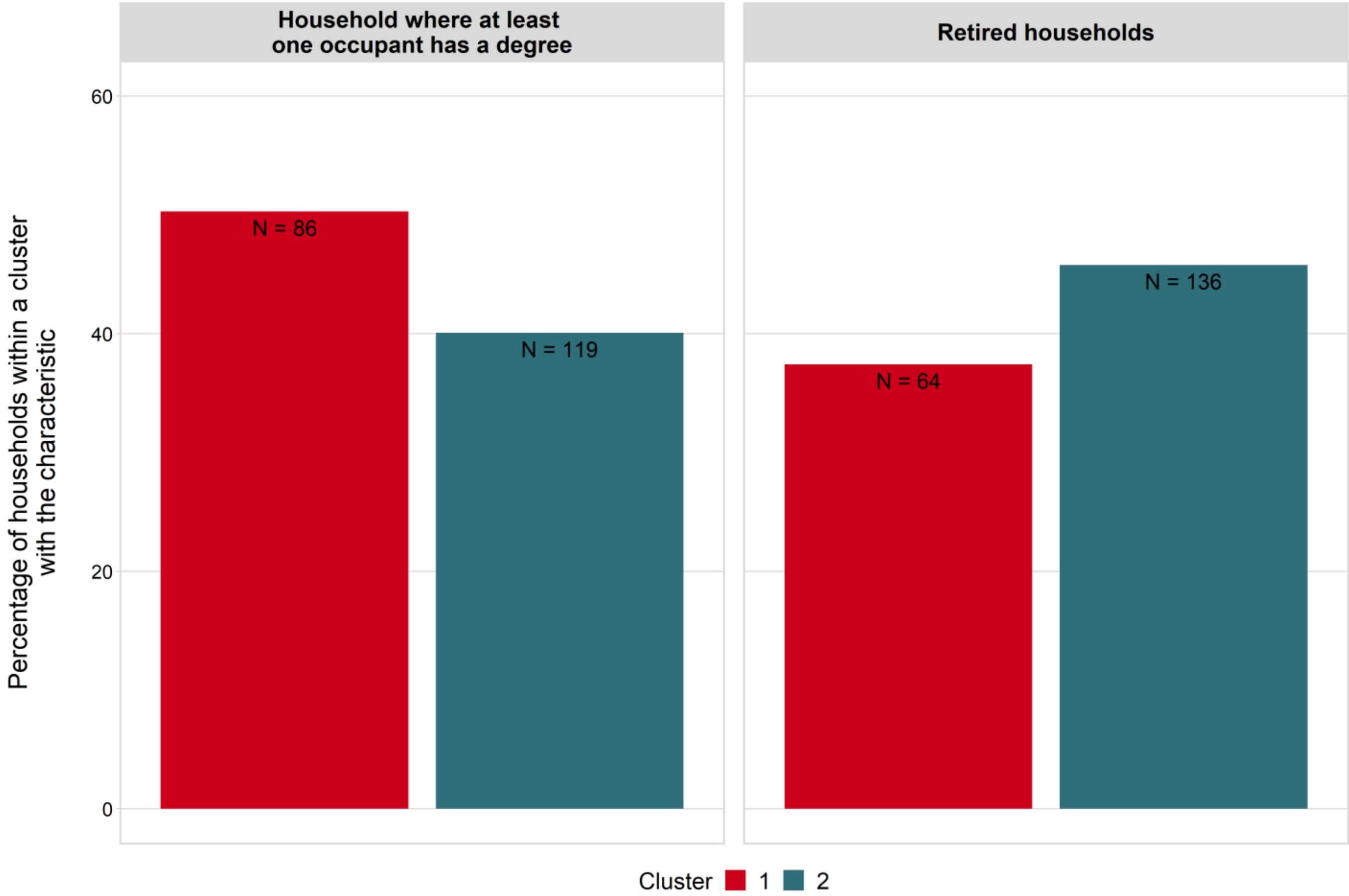


Source: Frontier Economics / UCL analysis using SERL data

The results are consistent with the pooled analysis in that they illustrate a substantial variation in responses. One cluster has an increase in both daytime and evening consumption across the sample, with the largest increases occurring during lockdowns. The other cluster has small changes in daytime consumption, some positive, some negative. Households in this cluster decrease evening consumption across the sample of months. The results are correlated over time – i.e. the households with the greatest response in one month or time of the day also tend to have a greater response in other months and times of the day.

We now consider the characteristics of households within these clusters. Figure 55**Error! Reference source not found.** below shows the characteristics of households within each cluster for the two cluster analysis.

Figure 55 Characteristics of households within each cluster for 2 cluster analysis

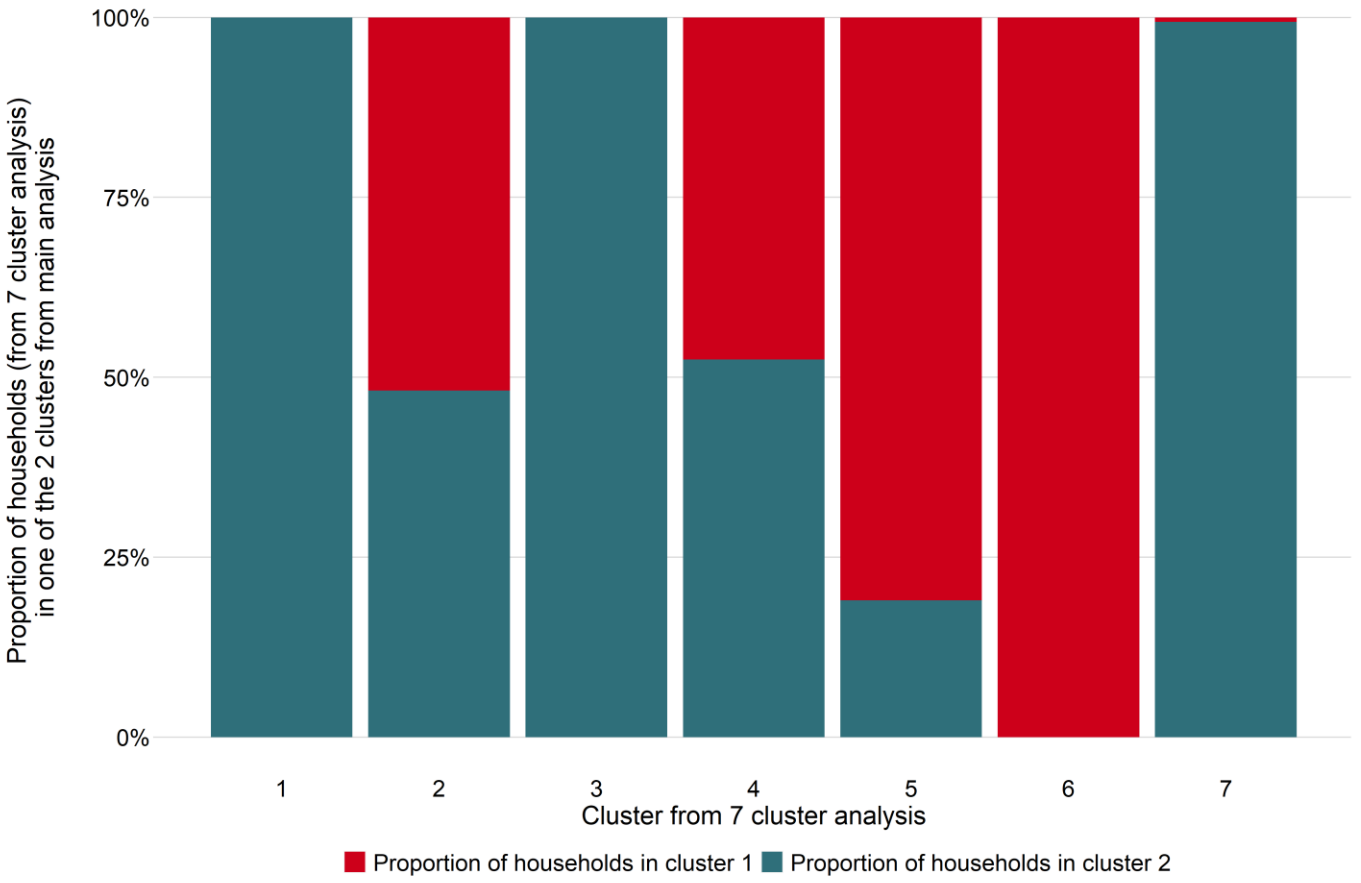


Source: Frontier Economics / UCL analysis using SERL data

In cluster 1 (the cluster exhibiting the greatest change in consumption), 50% of households have at least one occupant who has a degree (which equates to 86 households), whilst 37% of households are retired (which equates to 64 households). In cluster 2, 40% of households have at least one occupant who has a degree (which equates to 119 households), whilst 46% of households are retired (which equates to 136 households). This indicates that households in cluster 1, which has increases in daytime and evening consumption across the months, are more likely to have an occupant with a degree and less likely to have retired occupants when compared to cluster 2.

Finally, we consider how well households map from clusters in the seven cluster analysis to clusters in the two cluster analysis (see 3.3.2). Figure 56**Error! Reference source not found.** below shows the mapping.

Figure 56 Mapping of households from 7 cluster analysis to 2 cluster analysis



Source: Frontier Economics / UCL analysis using SERL data

The results show a reasonably clear mapping. Households from clusters 1, 3 and 7 are clearly in cluster 2 (of the 2-cluster analysis). There are 227 households in these three clusters, and they are 47.7% of the total sample of households. Households from clusters 5 and 6 are clearly in cluster 1 (of the 2-cluster analysis). Therefore around two-thirds of the There are 134 households in these two clusters, and they are 28.2% of the total sample of households. Clusters 2 and 4 are less clear-cut; the households are spread across the two clusters. However over three quarters of the households fall into one of the seven clusters which clearly maps to one of the two 'high response' and 'low response' clusters

The reasonably clear mapping suggests the seven-cluster analysis is a disaggregated version of the two-cluster analysis – i.e. the additional clusters are helping to break apart patterns of behaviour that the two cluster analysis shows.

Pooled months analysis

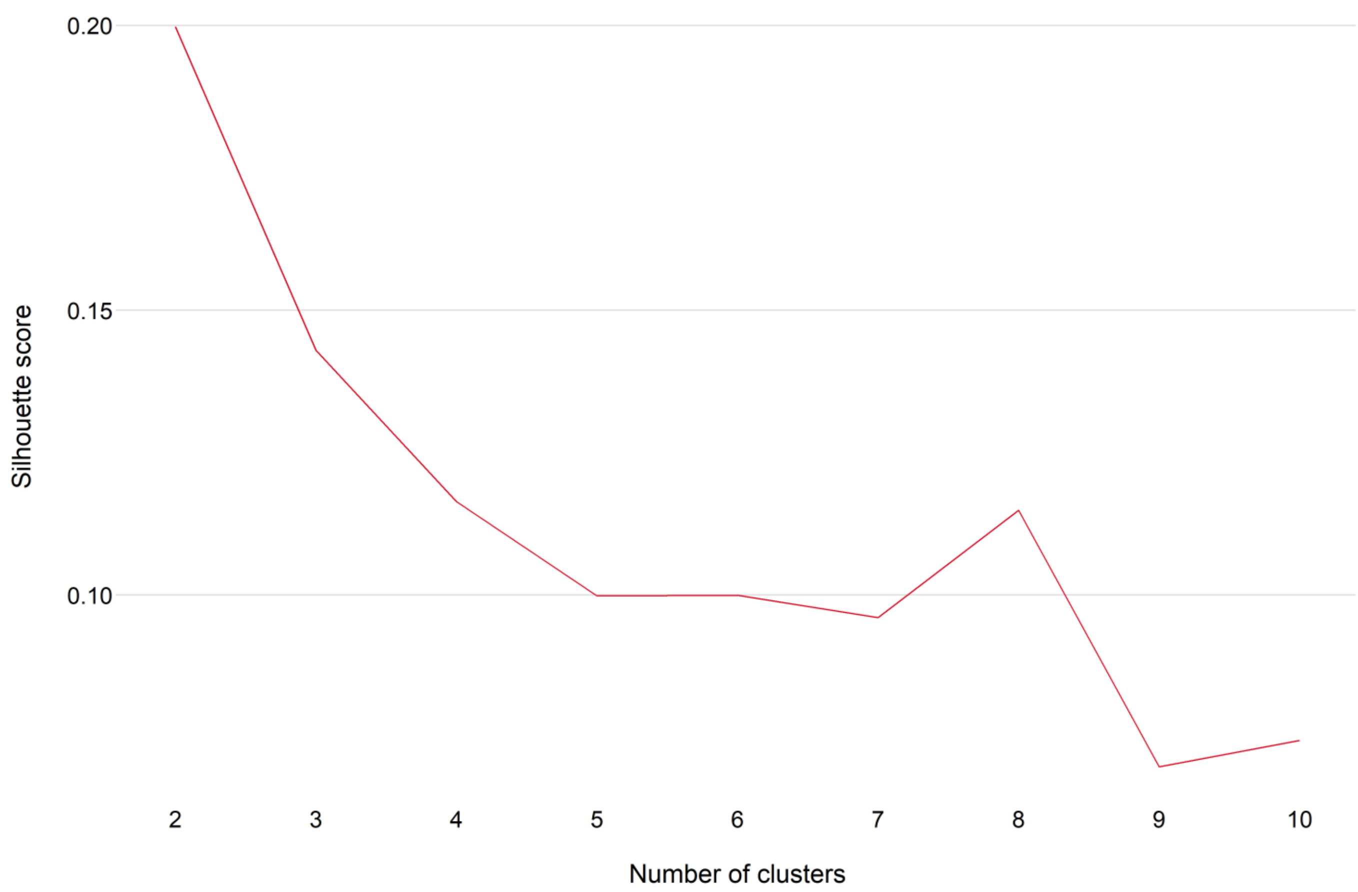
The analysis in the main body of the report separately considered customers' change in the peak time and evening, for six months (chosen to reflect lockdown and non-lockdown periods).

In this section, we describe the results of sensitivity where rather than using a single month of data, consumption changes have been pooled across two adjacent months. The use of more data can help ensure the results are less likely to be driven by random noise. However, given the short length of some of the lockdown and non-lockdown periods, we are only able to consider four pairs of months:

- April and May 2020 (lockdown)
- August and September 2020 (non-lockdown)
- February and March 2021 (lockdown)
- October and November 2021 (non-lockdown)

We now determine the optimal number of clusters. Figure 57 below shows the silhouette score for the pooled analysis.

Figure 57 Silhouette score pooled analysis

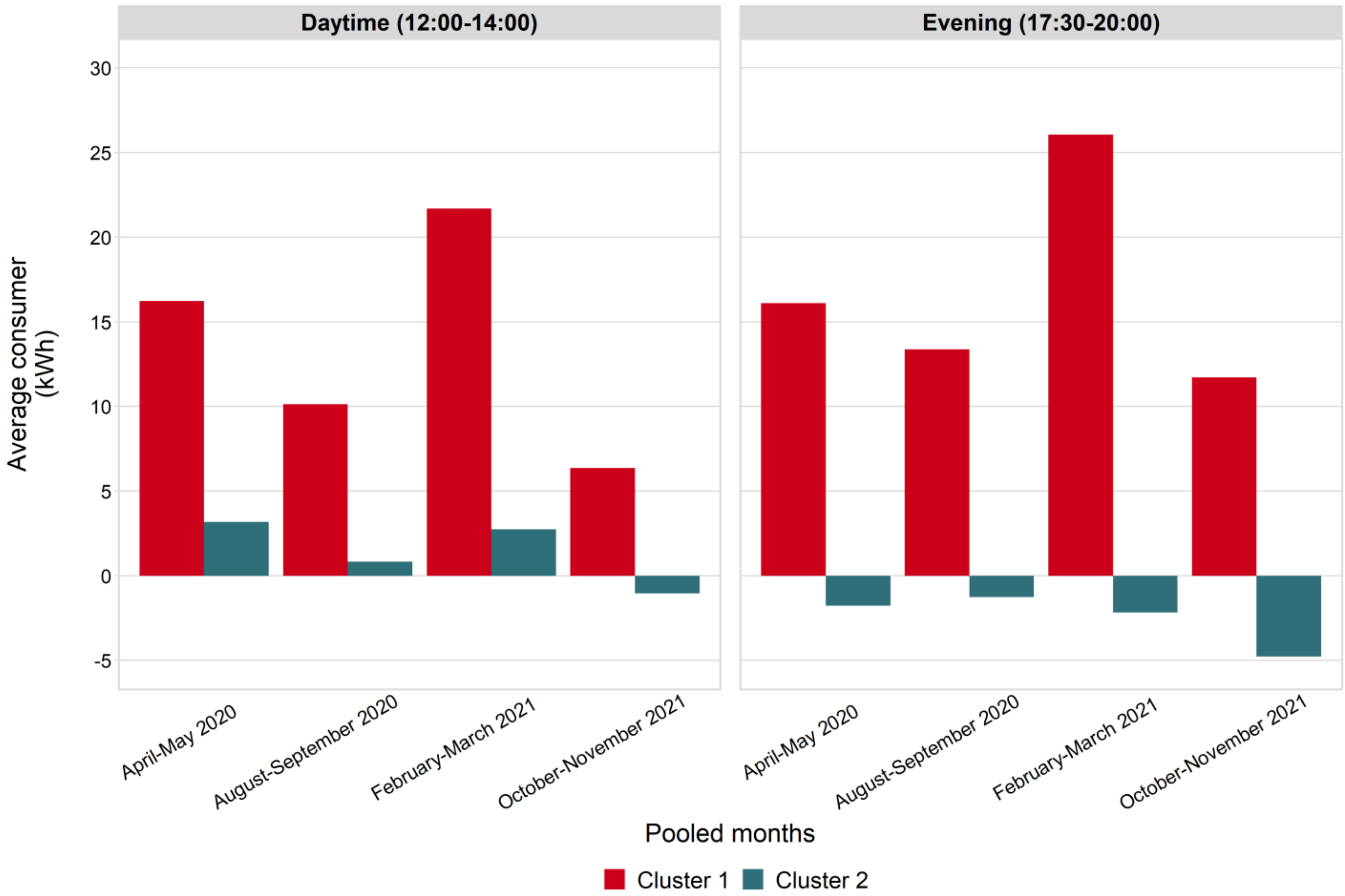


Source: Frontier Economics / UCL analysis using SERL data

As with our non-pooled analysis, the largest silhouette score occurs when the number of clusters is set to two, and so this is the number of clusters we choose.

Figure 58 below shows the change in consumption for the pooled analysis for the two clusters.

Figure 58 Change in daytime and evening consumption over time by cluster for pooled analysis



Source: Frontier Economics / UCL analysis using SERL data

As before, there are two distinct clusters. One cluster shows a positive change in both daytime and evening consumption, with the largest changes occurring during lockdowns, particularly the winter 2020/21 lockdown. The other cluster shows smaller changes in daytime consumption – for most of the pooled months the change is positive, whilst there is a decrease in consumption for October-November 2021. However, evening consumption in this cluster tends to be negative and increasing in magnitude as the pandemic period goes on.

Overall, the results are generally very similar to the version without pooled months.

Single month analysis

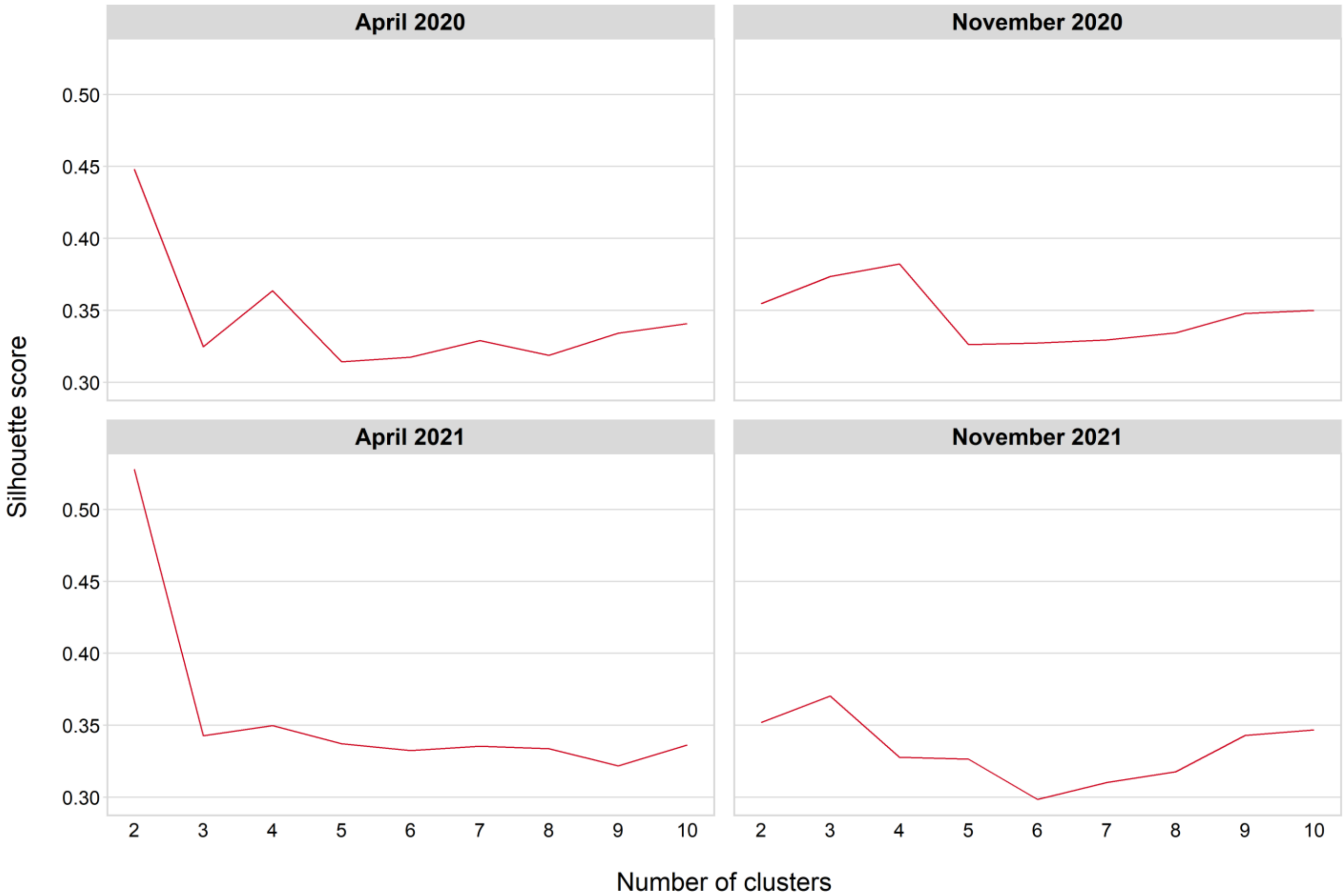
In this section, we describe the results of clustering analyses for individual months, and we also consider the characteristics of households within each of these clusters for the analyses. By focusing on a single month at a time, it is not possible to analyse how consumer behaviour persists across months (e.g. whether the same individuals with the greatest changes in consumption in April 2020 also had the greatest changes in April 2021). However as only two variables are used for each month (peak-time and evening consumption) it is easier to graph the results in a more intuitive way.

We first consider the clustering analysis for four given months, to understand how customers behave during different snapshots in time. We consider April 2020 (the first month of our sample, and a month characterised by a full lockdown) and November 2021 (the penultimate month of our sample, and a month characterised by no restrictions). We then choose two other months, November 2020 and April 2021[81], to understand how consumption changed within the sample.

We now determine the optimal number of clusters. Figure 59 below shows for silhouette score for each month.

[81] We have chosen November 2020 and April 2021 to ensure that months from across the sample are selected for the analysis. In addition, we have selected these months so that different types of restrictions are considered in the analysis: November 2020 was characterised by a full lockdown (as is April 2020 in the single month analysis), whilst April 2021 was characterised by a light lockdown.

Figure 59 Silhouette scores for April 2020, November 2020, April 2021, and November 2021



Source: Frontier Economics / UCL analysis using SERL data

For April 2020 and April 2021, the largest silhouette score occurs when the number of clusters is set to two, and so in the subsequent analysis we set the number of clusters to two. For November 2020 and November 2021, the largest silhouette score occurs when the number of clusters is set to four and three respectively, and so in the subsequent analysis we set the number of clusters to four for November 2020 and three for November 2021.

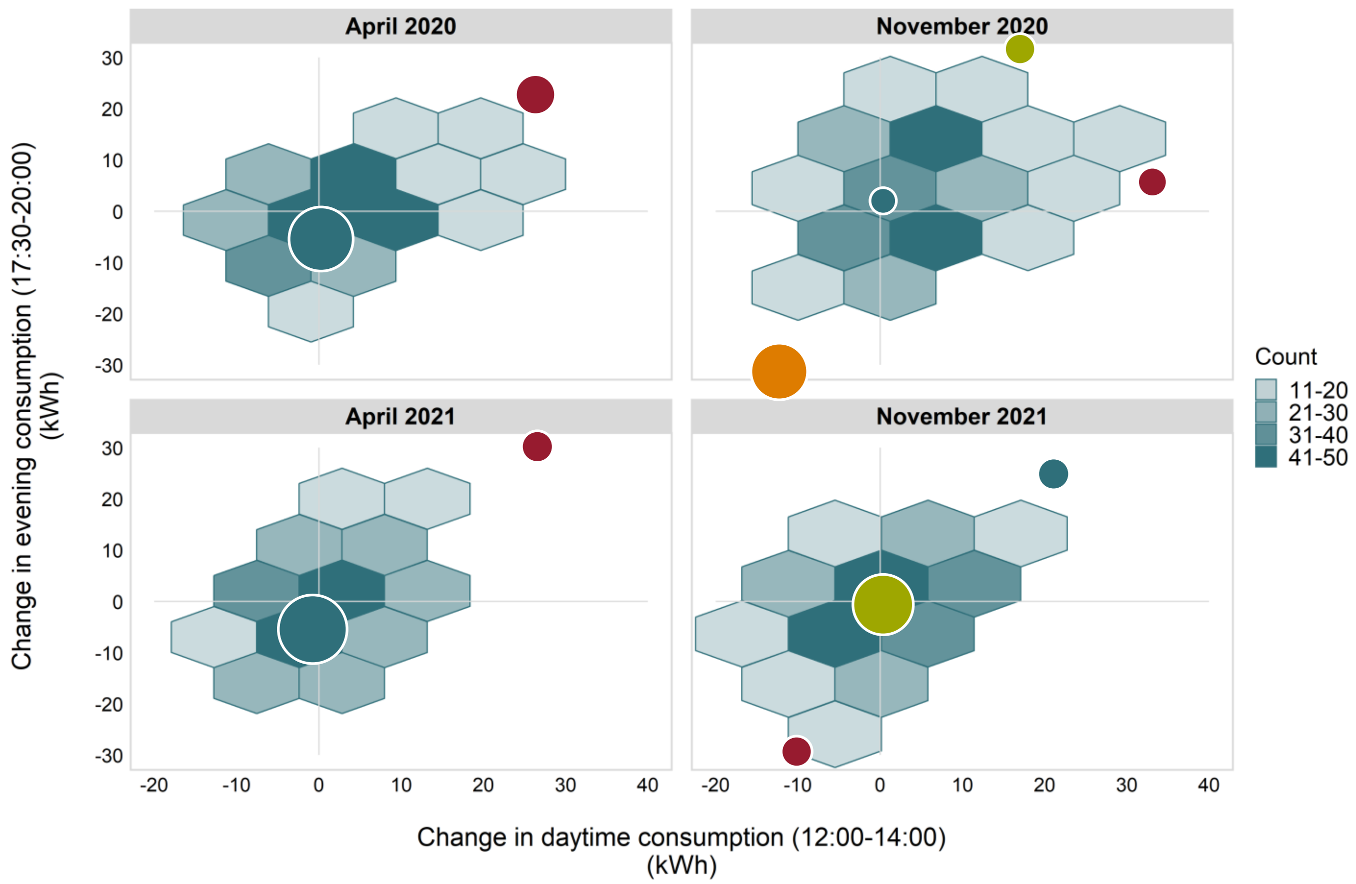
The single month analysis is two dimensional, as we consider the change in daytime consumption and change in evening consumption over one period of time. We can therefore visualise the distribution of changes in daytime and evening consumption, this is shown in Figure 60 on a hex plot. Within each hexagon, there are a number of observations, and the range of observations per hexagon is denoted by the colour of the hexagon. Note that we only show hexagons with more than ten observations to not disclose the consumption of individual households.

Approximate[82] positions of the cluster centroids (described in the following section) are also shown as coloured dots. The area of the dots is proportionate to the number of households in the cluster. In many cases, the cluster centroids are outside the hexagons (i.e. they relate to a pattern of consumption change

82 Centroids falling outside the domain of the graphs are shown on the edge.

that fewer than 10 households carried out, and is therefore redacted from the graph). This helps illustrate how widely varying the changes in consumption are, as well as demonstrating how cluster analysis can help visualise variations in smart meter data without revealing individual datapoints.

Figure 60 Hex plots showing the change in daytime and evening consumption



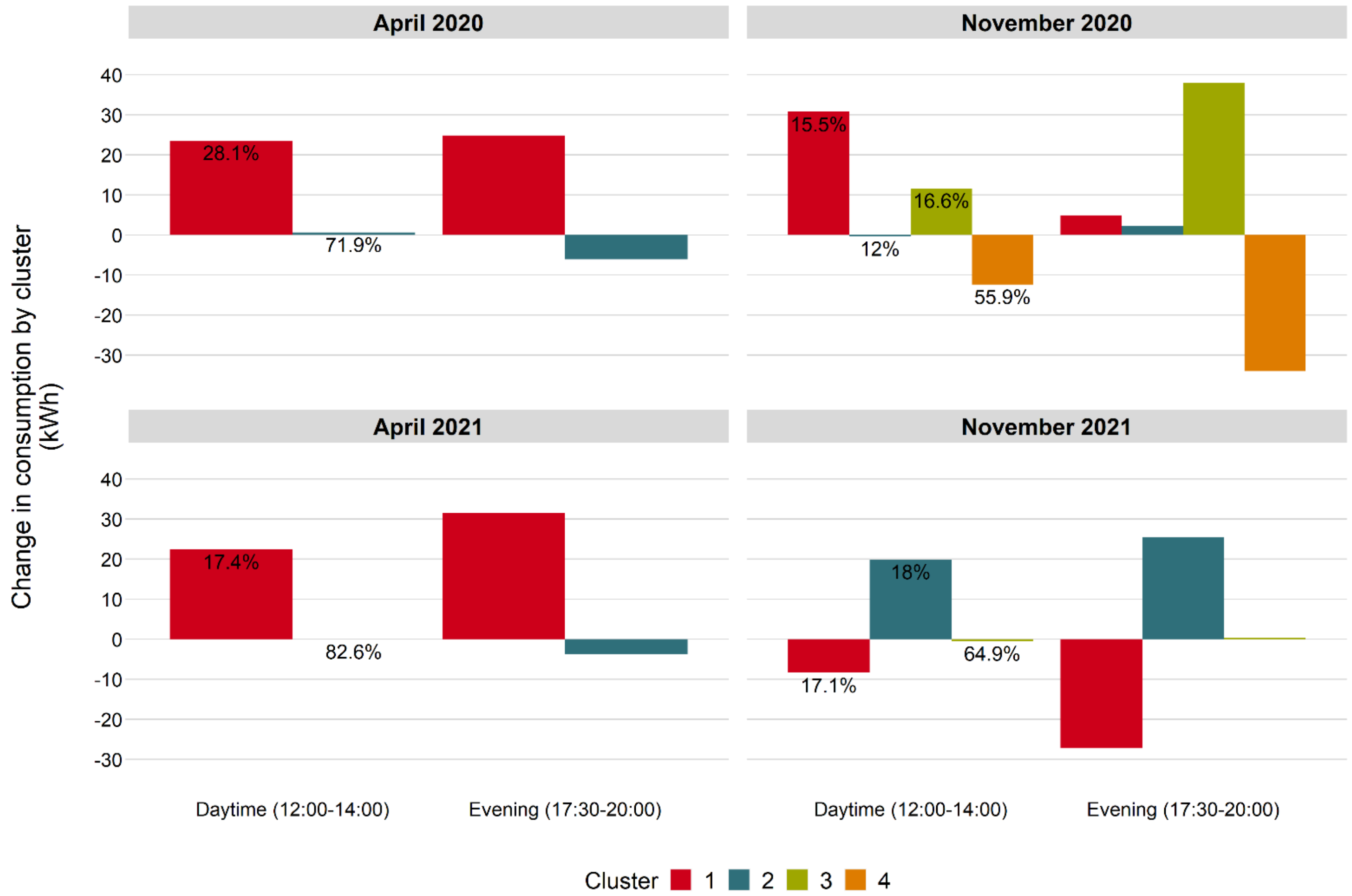
Source: Frontier Economics / UCL analysis using SERL data

Note: For context, 10 kWh is equal to an oven being used for around four and a half hours. See: https://www.daftlogic.com/information-appliance-power-consumption.htm

Figure 60 shows positive correlation between changes in daytime consumption and changes in evening consumption, and shows a slight skew towards positive changes in consumption, although there are some customers with negative changes in consumption, particularly in the November 2021 analysis. Most customers tend to be centred around no change in consumption, or a small increase in consumption. An exception is November 2020, where we know the optimal number of clusters is four, and we can see that there is a dense grouping with a slight negative evening change but positive change, whilst there is a dense grouping with both a positive daytime and evening change in consumption.

Having earlier determined the optimal number of clusters for each analysis, we can run the K-means clustering algorithm and obtain the average change in consumption for each cluster (which is equal to the value of the centroid for each cluster). Figure 61 below shows the different clusters for each of the selected months, and the change in consumption by cluster. The average change in consumption for a cluster is denoted by the height of the bar for that cluster. As each month is a separate analysis, households remain in the same cluster for daytime and evening consumption within a month, but can be in a different cluster across months.

Figure 61 Change in consumption by cluster in a given month



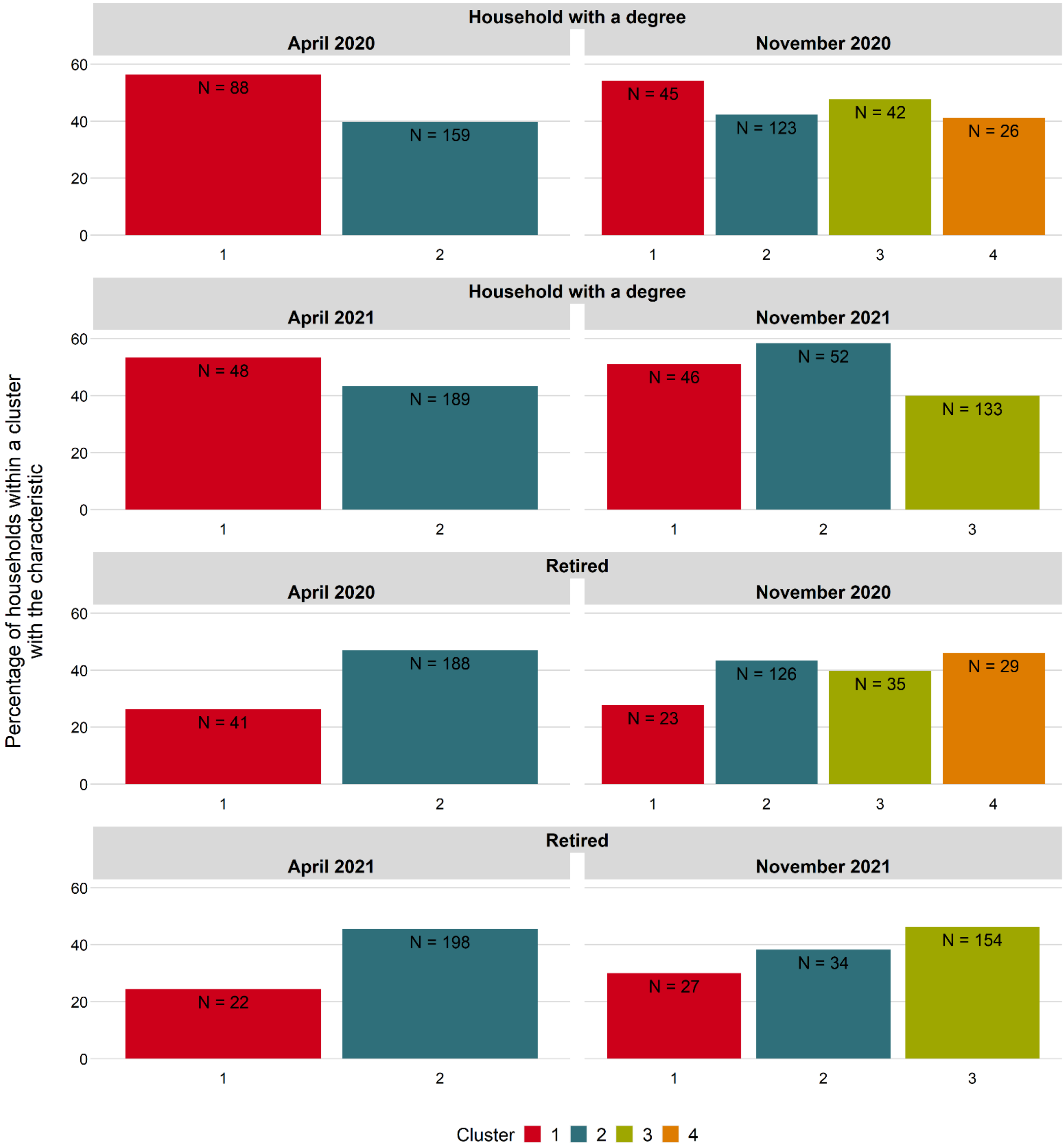
Source: Frontier Economics / UCL analysis using SERL data

Note: The percentages refer to the proportion of the total sample of households within each cluster, whilst the x-axis refers to the change in consumption for a cluster in either the daytime or evening

The results of the single month clustering analysis show there are clearly distinguishable groups of customers. From the April 2020 and April 2021 analysis there are two distinct groups: those who change consumption and those who tend to maintain existing consumption levels. For November 2020 and November 2021, there are a significant cluster of customers that tend to reduce consumption.

We next identify observable characteristics of households within these clusters, and this is shown in Figure 62.

Figure 62 Characteristics of households within each cluster for April 2020, November 2020, April 2021, and November 2021



Source: *Frontier Economics / UCL analysis using SERL data*

Note: *'N' denotes the absolute number of households within each cluster*

We calculate what proportion of each cluster is made of each type of household. This finds that households with retired members are associated with clusters showing no or negative change in consumption during the pandemic. In addition, households with degrees are associated with clusters showing positive changes in consumption.