Environmental Impact of Housing in England

Josh Cowley (Newcastle University)

May 2022

Contents

List of Figures

1. English Housing Survey (EHS)

1.1. Disclaimer

Data in this report referred to as the English Housing Survey (EHS) was obtained from the UK Data Archive (UKDA) under their End User Licence agreement.

As such, any data collection has been enacted for the purpose of non-profit research and further use of the data for commercial use without first obtaining permission from the data service provider is prohibited.

Furthermore, since these analyses can be interpreted as 'material derived from the data collections', any user of this report must also be in agreement to these terms.

1.2. Aims

The aim of this document is threefold:

1. reveal environmental-specific insights into the EHS dataset,
2. model the current EPC rating system based on the factors that can exist in a dwelling,
3. understand regional trends in dwelling efficiency while accounting for national trends.

By combining these main aims, we will be able to create an environmental indicator that serves as a summary to the impact of housing on the climate.

To achieve this, we explore features of the EHS dataset that are pertinent to the energy efficiency and CO2 emissions of the sampled dwellings. Then, we employ regression to better understand the relationships between the ever-adapting EPC ratings and physical attributes of dwellings such as insulation and type of boiler installed.

The final aim is realised using a multi-level modelling framework that allows practitioners to separate population-level effects with group-level effects where the population in this context is at a national level and the regions, such as the North East, are groupings. It is these parameters and subsequent estimates that are intended to serve as environmental indicator(s).

1.3. Sampling Methodology

The data collected for use in this report are available, with full documentation, at the UK Data Archive (Series 200010). Documentation compiled by the EHS relay, in detail, how the data was collected.

In summary, an initial sample of addresses is pulled from the Postcode Address File (PAF) using a two-stage random sample that is not covered in any detail here, for more information see EHS Technical Reports. Participants from this initial sample are invited to take part in an *interview* stage survey that is based on **households**. A subset of these participants where permission is granted or the property is vacant are then subject to a *physical* stage survey. As such, this 'opt-in' approach may lead to a sample bias that is yet to be considered.

Note the distinction of terms:

- **household(s)** refer to the group of people (or person) who have the accommodation as their only or main residence,
- **dwelling(s)** refer to the unit of accommodation which households may inhabit.

This report concerns itself with dwellings only and does not utilise data from the initial interview survey. Further work should consider this extension to the dataset as a means of understanding fuel poverty and the household-dwelling relationships.

Figure 1.1 shows unequal, yet consistent, sample sizes across regions. This is expected as the number of houses in each region are not necessarily equal.

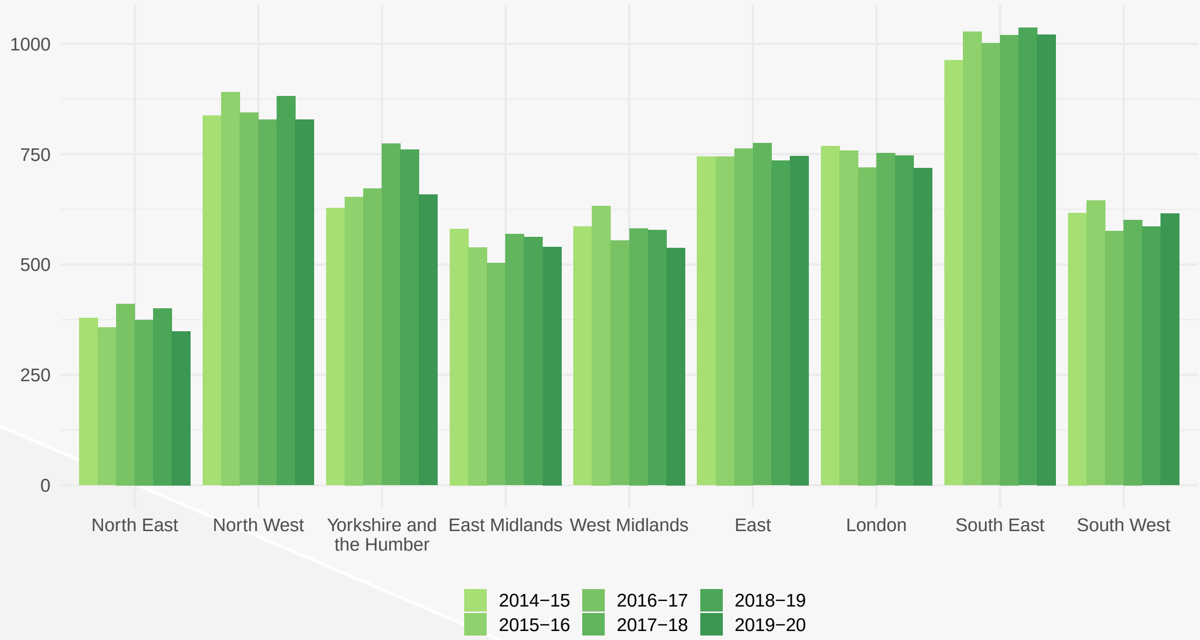

Figure 1.1.: Unequal sample size across regions

1.4. Dwelling Stock

1.4.1. General

Here we are interested in the general properties sampled in the EHS, the reader is also advises to view the annual headline report (Department for Levelling Up, Housing and Communities, 2020a). For example, Figure 1.1 in the 2020-21 version shows an increase in private renters and a decrease in social renters from the start of the millennium.

Figure 1.2 shows us the proportion of properties that are owned by *housing assocations*. Both the North East and North West have a higher proportion when compared to other regions with no minor signs of change. This could be due to a number of reasons such as fewer renters (when compared to London) or fewer households being able to get on the property ladder.

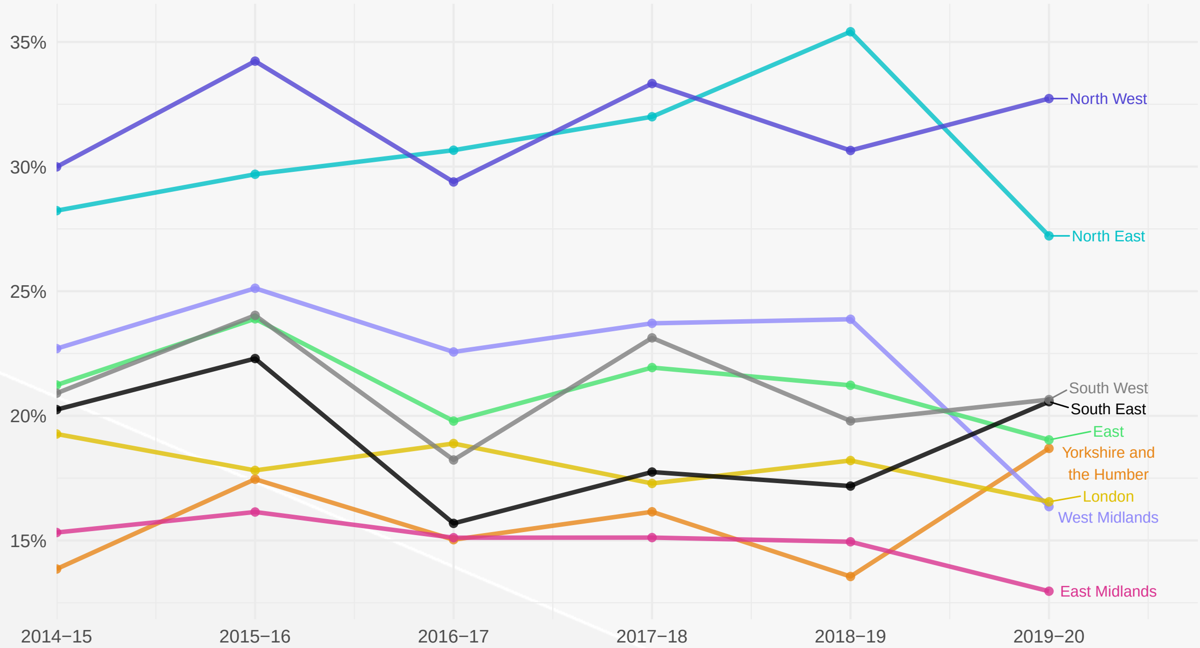

Figure 1.2.: Housing association market share highest in Northern regions

Properties that are vacant are few and far between in the data; all regions report 96% to 98% of sampled properties are inhabited for the period between April 2014 and March 2020.

Each dwelling is assigned into one of 10 categories based on the type of dwelling and when it was built. One clear discrepancy in the regions that is highlighted by these data is the proportion of flats (defined as a purpose built flat or converted flat) as shown in Table 1.1.

Table 1.1.: Proportion of flats by region

Region	Number of Flats	Proportion
London	2581	57.8%
South East	1549	25.5%
South West	819	22.5%
East	935	20.7%
North West	1033	20.2%
West Midlands	683	19.7%
Yorkshire and the Humber	775	18.7%
North East	418	18.4%
East Midlands	504	15.3%

Note:

EHS Physical Sample (April 2014 to March 2020)

Also reported in the EHS is if there is any damp present, damp could be a symptom of bigger problems. Thankfully, the percentage at which this occurs is down from 5.18% in 2014-15 to 4.01% in 2019-20 although to understand if these changes are indicative of an actual decline or just variance in the data would require further investigation over a larger timescale.

1.4.2. Energy Specific

1.4.2.1. Space Heating

Figure 1.3 shows us that there is a clear reliance on gas for systems related to heating up space with no clear trend emerging from the data. In Department for Business, Energy & Industrial Strategy (2019), it is claimed that significant increases in heating bills and lifetime costs may pose a barrier that would explain this hesitancy to switch to electrical systems.

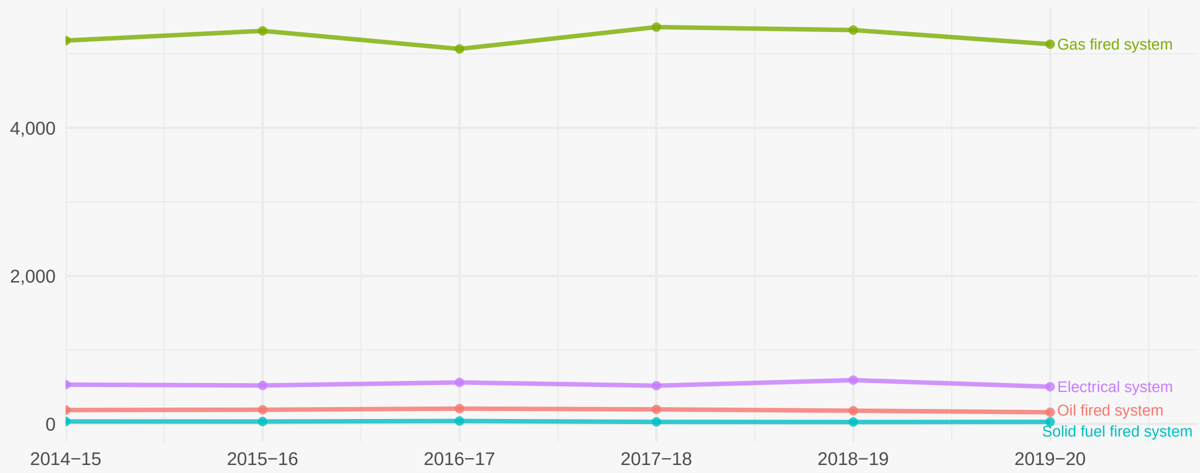

Figure 1.3.: Little to no change in heating systems

1.4.2.2. Wall Types

Wall type data refers to the predominant external wall of a dwelling. If a surveyor deems at least 50% of the external wall to be of cavity or solid construction, that is recorded, otherwise it is recorded as "other". Similarly, if a surveyor deems at least 50% of the wall to be insulated, that is also recorded in the data. As such, any dwelling can be one of 5 types summarised in Table 1.2

Note that the insulation frequency of cavity walls is much higher than the solid walls. As expected due to the difficulties involved with insulating a solid wall relative to a cavity type wall.

Table 1.2.: Insulation more dominant in 'Cavity' type walls

Wall Type	Uninsulated	Insulated	Unknown
Solid	8,404	1,399	N/A
Cavity	8,090	18,196	N/A
Other	N/A	N/A	882

1.4.2.3. Boiler Type

For any dwelling that uses central heating, we have information on the type of boiler used. If we were to look across the regions we see a similar trend occurring that is summarised in Figure 1.4. That is, the volume condensing boilers in each sample is increasing with time.

British Gas (2022) claim that condensing boilers these are generally the most efficient boilers "on the market" and that "nearly all modern boilers are A-rated condensing boilers" which is a good sign of newer technology that is better for the environment is being implemented across all regions.

As with any emerging technology, it is important to remember that a minor increase in efficiency may not justify a full boiler replacement and the oldest or least efficient appliances should be replaced as a priority.

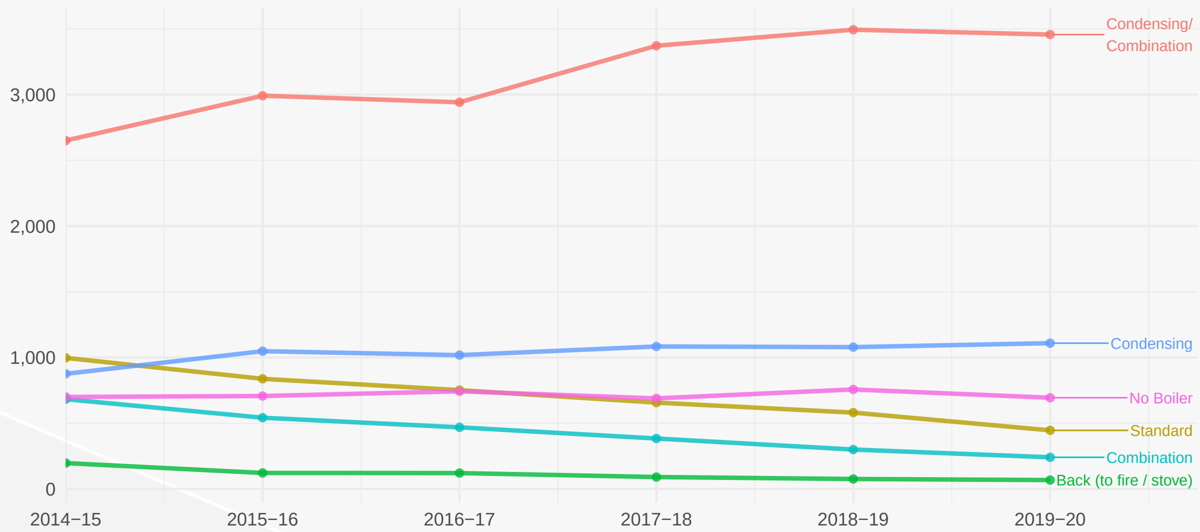

Figure 1.4.: Newer boiler types being adopted, nationwide

1.4.3. Energy Ratings

Although the EHS provide two metrics that measure different attributes, that is, the energy efficiency rating measures energy efficiency and the energy impact rating measures CO_2 emissions, one would expect these to be very similar.

In fact, in our subset of the EHS sample we observe that the bands for each metric are equal 66.4% of the time and are at most 1 band different in 98.8% of cases.

As such, Figure 1.5 shows only the efficiency ratings, but due to high correlation, can be viewed as representative of both ratings. It is clear that the majority of dwellings are in the middling bands of C and D, specifically 83.8% of our data. There are so few being in the top 2 bands of A and B that the surveyors grouped them together in the data.

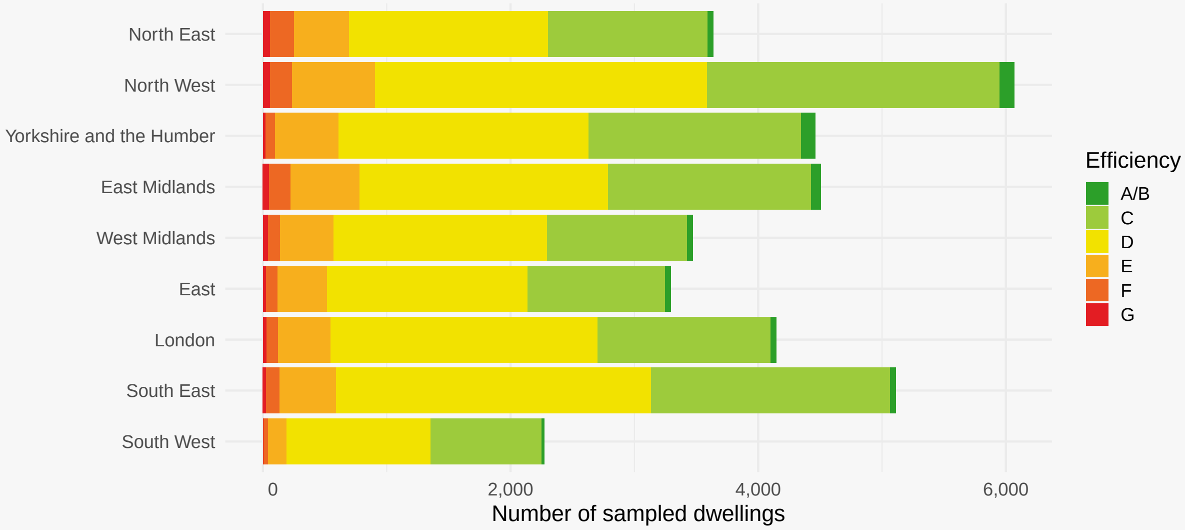

Figure 1.5.: Most dwellings categorised as C or D bands (EPC)

While these bands are a good simplification for consumers, we aim to model the underlying Standard Assessment Procedure (SAP) ratings that have lead to these groupings. Figure 1.6 shows a histogram of all SAP ratings in our data and colours them according to their bands; note that the distribution is left-skew meaning we expect to see more extremely inefficient houses over extremely efficient ones. There are also several outliers with a SAP rating reported as 1, further investigation into these outliers is required.

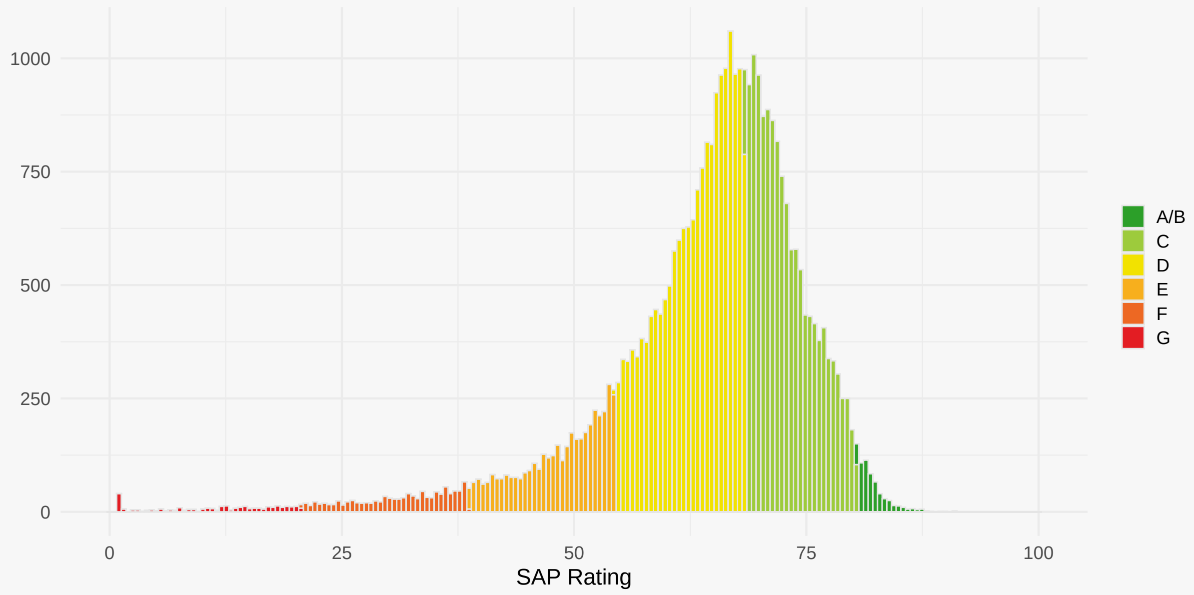

Figure 1.6.: Efficiency ratings average around 64 with a left-skew

1.5. Data Availability

1.5.1. EPC

Energy and impact ratings are measures that are considered to be indicators of efficiency and emissions respectively. However, these ratings are subject to undergoing changes in an effort to improve their accuracy and utility to decision makers. For example, there was a *call for evidence on EPCs* (Department for Levelling Up, Housing and Communities, July 2018) that aimed to evaluate the current system and improve any shortcomings.

Hence, in an ideal scenario we would have data on energy used, energy demand and CO2 emissions per house, however such an undertaking on a dataset of this scale is simply infeasible and unlikely to be improvement on the already scrutinised EPC measure.

1.5.2. Renewable Technologies

Omitted from the EHS is the mention of renewable technologies (RT). Such technologies are a significant investment and as such will have a substantial impact on a households energy bill and also the energy efficiency rating.

Datasets such as the Local Authority Housing Data (Department for Levelling Up, Housing and Communities, 2020b) contain data on the total number of RT(s) a local authority has installed in its own dwellings that are leased to social renters. These data also indicate which RT(s) such as heat pumps or biomass boilers but neglect to reveal the degree of which each is utilised.

1.5.3. Retrofit Impact

As a product of previous improvements, many dwellings are in a condition where improvements to efficiency or emissions are more ‘difficult’. That is, a higher economical cost, more resources need or higher skilled labour. For example, a dwelling with mainly cavity walls is much more likely to be insulated than a dwelling with solid walls for this reason, see Table 1.2.

Data on the environmental cost of such retrofits are meagre and no such dataset has been found during the research phase of this project.

To achieve the carbon neutral goals set out by the UK government and others in the Paris agreement, we must ask is the carbon investment and other pollution costs worth the offset that they will achieve. Similar to how various improvements are financially infeasible due to cost, we must collect more data to rule out improvements that have too high an environmental cost.

2. Model A

2.1. Modelling EPC

An EPC certificate contains various information that is then used to calculate two typically correlated measures:

1. Energy efficiency rating (EER).
2. Energy impact rating (EIR).

These ratings range from 1 to 100 where higher values indicate dwellings that are better for the environment. Data is collected using a Standard Assessment Procedure (SAP) in an attempt to standardise these values across heterogeneous properties.

The energy efficiency rating measures the efficiency of the entire building and is framed to a tenant as a key influence on running costs. That is, higher values of EER can not only yield benefits for the environment (as less energy will be used) but also for the household paying the energy bill.

As shown by the sample certificate provided by the UK Government, the EIR is more indicative of CO2 emissions (UK Government, 2022).

> The environmental impact rating is a measure of a home's impact on the environment in terms of carbon dioxide (CO2) emissions. The higher the rating the less impact it has on the environment.

Expectedly, these two ratings tend to be correlated; in our cohort of dwellings the Pearson correlation coefficient is as high as 90.5%.

As previously mentioned, the methodology that creates both ratings has been under some criticism (Department for Levelling Up, Housing and Communities, July 2018) and hence any measure derived from these data will inherit any limitations.

2.1.1. Model Description

Our initial model comprises of two multivariate linear regressions, one for each rating described above. The aim of such models is to verify that the variables that experts would think are important to energy efficiency and emissions are then also highly correlated with EER and EIR respectively.

Currently we make use of 6 predictors in our model.

- boiler type,
- double glazing percentage,
- fuel used for space heating,
- degree of loft insulation,
- degree of wall insulation and wall type,
- system used for water heating

We expect all predictors will be statistically significant due to the data source, we reason that these specific variables are collected because they are seen as important to EPC ratings, and as such this section is a proof of concept for a greater set of variables.

Greater benefits would be realised from using predictors that are less prominent, such as renewable technologies highlighted in the data availability section, where parameter estimates would potentially highlight a shortcoming in the current rating system.

Table 1.2 shows an estimates for various categories found in the data. One can calculate an expected EER and EIR by finding the corresponding attributes and summing the estimates (the intercept is added for all dwellings).

This leads to the intuition that we can compare attributes to each reference level and infer its impact on the rating, subject to keeping all other attributes fixed.

For example, see that the reference level for boiler type is "Standard" and the estimate for "Combination" boiler is between 2.6 to 3.3 with a significant P value ($p < 0.05$). Thus, we expect that upgrading a house from a standard to a combination boiler would increase their EER by approximately 3 points.

2.1.2. Discussion

Table 1.2 shows us that a dwelling with a standard boiler, no double glazing or insulation that is still reliant on oil systems and an electric water heater would expect to have a G band EPC rating. This hypothetical does not meet the current minimum energy efficiency standard (MEES) allowed for rented properties and such an anti-environment dwelling is likely in need of more pressing repairs related to safety and could even be vacant.

The impact of different types of boilers is typical with the worst being a back boiler, as expected. Interestingly, the impact on efficiency for condensing and condensing combination boilers appear to be very similar but the latter appears to produce less CO2 emissions on average.

Double glazing in many modern homes is a necessity and can be seen as a good candidate for retrofitting, the data supports this implying a upgrade of *all* windows in a house may increase the EER by as many as 7 points.

For space heating there is a single high rating choice, gas fired systems. There is not significant evidence to distinguish the other options. As mentioned in an earlier section (see Figure 1.3) electric heating systems have higher running costs and the efficiency comparison needs to be investigated further. A similar story is told by the water heating estimates, although it should be noted that a dedicated boiler could include both gas and electric boilers.

Loft insulation appears to increase both EER and EIR as expected but as the difference between estimates decrease it may be a case of diminishing returns and warrants further investigation.

Wall insulation is less straightforward to compare since the data has encoded both insulated and wall

Table 2.1.: Linear Regression Results

	Efficiency Rating		Impact Rating	
Term	Estimate (95% CI)	P Value	Estimate (95% CI)	P Value
(Intercept)	14.6 (13.0 to 16.3)	<2e-16	5.3 (3.3 to 7.2)	1e-07
Boiler				
Standard	Reference		Reference	
Back (to fire / stove)	-4.8 (-5.4 to -4.3)	<2e-16	-7.3 (-8.0 to -6.7)	<2e-16
Combination	3.0 (2.6 to 3.3)	<2e-16	4.1 (3.7 to 4.5)	<2e-16
Condensing	7.7 (7.5 to 8.0)	<2e-16	8.6 (8.3 to 8.9)	<2e-16
Condensing-combination	7.7 (7.5 to 8.0)	<2e-16	10.1 (9.9 to 10.4)	<2e-16
Double Glazing				
No double glazing	Reference		Reference	
Less than half	1.2 (0.6 to 1.9)	2e-04	-1.1 (-1.9 to -0.3)	0.005
More than half	3.4 (2.9 to 4.0)	<2e-16	1.8 (1.2 to 2.5)	2e-08
Entire house	6.5 (6.1 to 7.0)	<2e-16	6.8 (6.2 to 7.3)	<2e-16
Space Heating				
Oil fired system	Reference		Reference	
Gas fired system	11.7 (11.3 to 12.1)	<2e-16	14.9 (14.4 to 15.4)	<2e-16
Solid fuel fired system	0.2 (-1.0 to 1.3)	0.789	1.9 (0.6 to 3.3)	0.005
Electrical system	0.6 (-1.0 to 2.2)	0.460	2.3 (0.4 to 4.3)	0.019
Loft Insulation				
None	Reference		Reference	
Less than 50mm	6.5 (5.7 to 7.2)	<2e-16	7.0 (6.1 to 7.8)	<2e-16
50 up to 99mm	9.4 (9.0 to 9.9)	<2e-16	10.1 (9.6 to 10.7)	<2e-16
100 up to 149mm	10.8 (10.3 to 11.3)	<2e-16	11.8 (11.2 to 12.3)	<2e-16
150 up to 199mm	11.4 (10.9 to 11.9)	<2e-16	12.8 (12.2 to 13.3)	<2e-16
200mm or more	12.2 (11.7 to 12.6)	<2e-16	14.0 (13.4 to 14.5)	<2e-16
Wall Insulation				
Solid uninsulated	Reference		Reference	
Cavity with insulation	9.5 (9.3 to 9.7)	<2e-16	12.2 (12.0 to 12.5)	<2e-16
Cavity uninsulated	5.1 (4.8 to 5.3)	<2e-16	6.2 (6.0 to 6.5)	<2e-16
Solid with insulation	9.1 (8.6 to 9.5)	<2e-16	12.7 (12.1 to 13.2)	<2e-16
Other	14.6 (14.1 to 15.2)	<2e-16	19.2 (18.6 to 19.9)	<2e-16
Water Heating				
Electric immersion heater	Reference		Reference	
With central heating	8.7 (7.2 to 10.2)	<2e-16	6.4 (4.6 to 8.1)	3e-12
Dedicated boiler	5.4 (1.8 to 9.0)	0.003	2.6 (-1.6 to 6.9)	0.228
Instantaneous	5.8 (3.4 to 8.3)	3e-06	6.7 (3.8 to 9.6)	5e-06

type. We have chose the solid uninsulated wall as a reference level although modifying walls between solid and cavity is clearly an infeasibility and only serves as a comparison on the population scale.

2.1.3. Further Work

As previously mentioned, this model could be improved by increasing the range of data available, as is the case for almost all models.

One prospect for this model could be decision makers in organisations holding dwelling stock as it could serve as a predictive model that highlights any properties that have rating less than expected and would benefit from the investment of a renewed EPC certificate.

Further work on the current data could be an investigation into interaction terms, the process of understanding the role of each predictor without the restrictive assumption of keeping everything constant. Leading to answers from questions such as “what is the impact of double glazing in a well insulated house compared to a less insulated house?”.

3. Model B

3.1. Modelling Regional Impact

3.1.1. Multilevel Modelling

Multilevel modelling allows us to distinguish between population effects and group level effects (Raudenbush and Bryk, 2002). In our context, the population would be all dwellings within England and each group can be asserted to be the regions. Trivially the scope can be decreased, say to a local authority level, or even increased, say to a worldwide level, depending on the application and available data.

In our model we fit intercepts only, namely:

1. a population effect for each year that the data spans,
2. a group-level effect for each year that the data spans.

Suppose n_y denotes the number of years our dataset spans; for the EHS data from 2014 to 2019, $n_y = 6$. This model will result in n_y population parameters to estimate, each summarising the SAP rating, irrespective of regions. Then, we also have n_y group-level parameters for each defined region that summarises how the SAP rating for that region compares to the overall population on a year-by-year basis. A formal description of this model is available in Appendix A for the statistically minded reader.

By inferring these parameters we can form an environmental indicator on a national level and a more granular indicator for each region that is derived from the energy efficiency ratings of the English

Housing Survey sample. We focus on efficiency but the methodology can be re-executed for the impact rating to shift focus from efficiency to CO2 emissions.

3.1.2. Bayesian Inference

To fit this model we make use of R and the popular brms package (Bürkner, 2021). R version: 4.0.3 and brms version: 2.16.3 is used. The aforementioned software will fit our model in Stan (Carpenter et al., 2017) that uses Bayesian methods such as Hamiltonian Monte Carlo and a No U-Turn sampler, neither of which will be explained in any detail here.

More importantly, these methods allow us to perform Bayesian inference in an efficient manner. When using the Bayesian paradigm we combine expert knowledge, known as a *prior*, with what the data tells us, a *likelihood*, to form our *posterior* beliefs.

In this report we use naive priors in the absence of clear expert opinion, should such an expert be made available for the benefit of this model, there exists many methodologies of prior elicitation that will increase the accuracy and efficiency of our model, for example see Gosling (2018).

3.1.3. Results

We run our MCMC scheme over 2 chains for 4,000 iterations (1,000 discarded as warm-up). Bürkner (2017) remarks that we need much fewer samples than a traditional MCMC scheme due to the higher quality of samples produced, that is, a higher effective sample size per iteration.

Figure 3.1 shows us the population estimates by year with uncertainty quantified by the 95% credible intervals of the parameters. Using knowledge of the EPC bands, we see that the estimate is representative of a D rating which is not overly surprising given that the majority of dwellings have a C-D rating. Positively, we see for the years we have selected the posterior estimate is increasing indicating an increase in efficiency ratings.

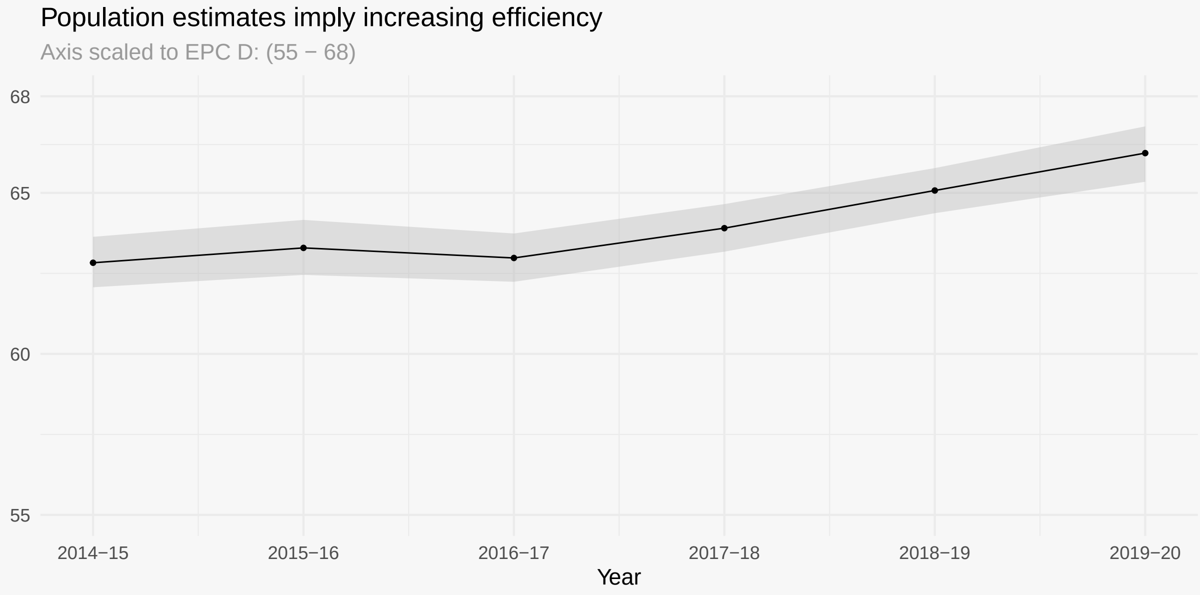

Figure 3.1.: Population estimates imply increasing efficiency

For a region-specific indicator one could combine the estimates from Figure 3.1 with those from Figure 3.2. However just by investigating the latter Figure, we see that no region's estimate(s) are substantially different to 0, implying no region is substantially different from the population. North east credible regions do exceed 0 implying a marginal statistical significance but only by a few points at best.

In the author's opinion, small differences on a regional level are likely due to the fact that substantial change in housing efficiency tend to come from *green* government scheme(s) that are available nationwide. Such a claim warrants further investigation and it is important to remember the models highlight correlation not causation.

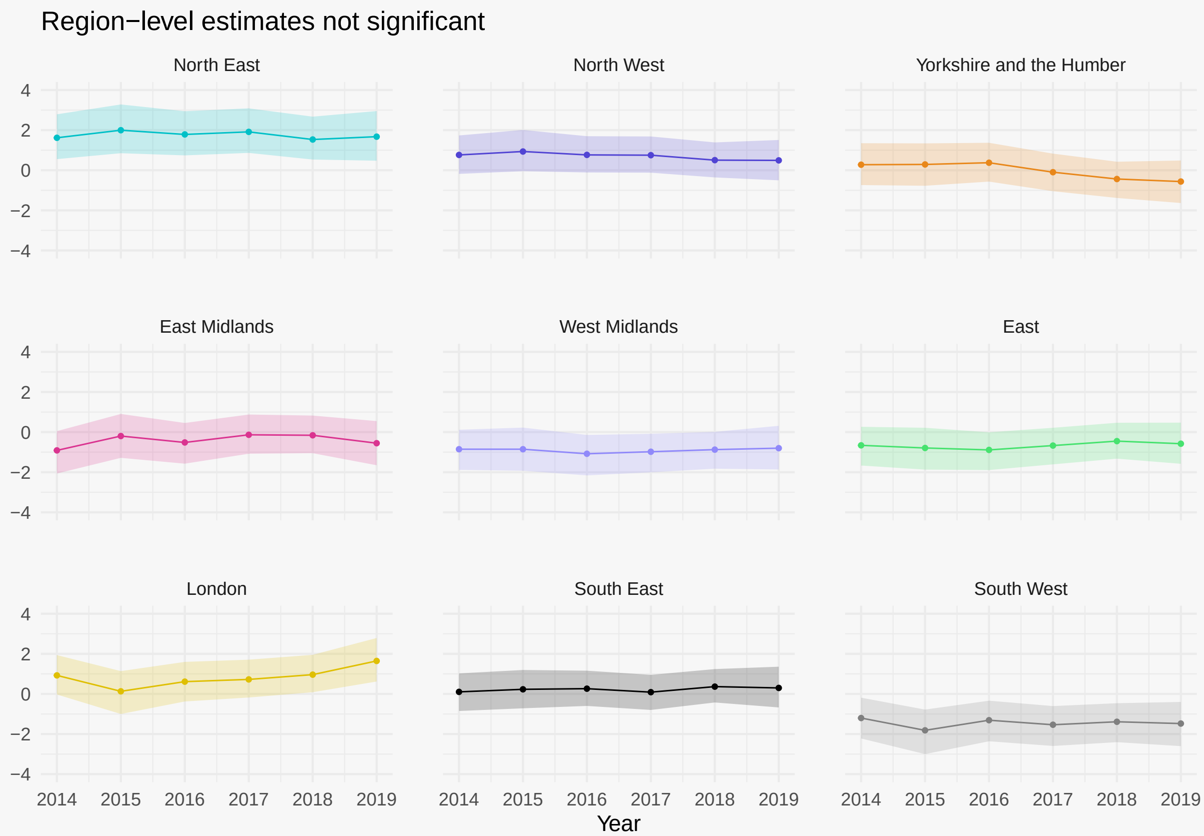

Figure 3.2.: Region-level estimates not significant

3.1.3.1. Noise Term

For both models we have assumed an independently and identically normally distributed error term that is required to allow for variance in the data. In this model, we summarise our estimate in the same way as the other parameters, by looking at the mean and credible region of the samples arising from the Markov chain Monte Carlo samples.

Table 3.1.: Error term summary

Term	Estimate	Lower 95% CI	Upper 95% CI
Sigma	11.14	11.06	11.22

The parameter summarised in Table 3.1 is trying to estimate is representative of the measurement error in the SAP rating and would ideally be minimised. A further investigation could use this variable as supplementary information to determine if the accuracy of *Standard Assessment Procedure*.

3.1.4. Model Extensibility

This model, in its current form is closely related to an intercept only regression and as such aims to describe data without any ambition to understand any underlying correlation.

It is straightforward to combine this model described in Chapter 2 to create a framework that is able to understand the impact of various predictors such as boiler type, insulation and others. However, unlike before we would be able to understand correlations not only at a population level but at a regional level as well.

Appendix

A. Multilevel Model Description

Suppose we let y_i be the variable we are interested in, such as energy efficiency rating, for a single year. Each observation or dwelling is indexed by $i = 1, \ldots, n$, where n is the sample size.

In a multivariate linear regression (model A), we leverage dwelling information such as insulation or boiler type. Denote these predictors by a p length vector $\boldsymbol{x}_i = \begin{pmatrix} 1 & x_{i1} & \ldots & x_{ip} \end{pmatrix}$ where the 1 is added to all data to fit an intercept (population average). We can now model the response by

$$y_i = \boldsymbol{x}_i^T \boldsymbol{\beta} + \epsilon_i,$$

where $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$. Here, $\boldsymbol{\beta}$ and σ are parameters to be estimated, specifically regression and noise parameters respectively.

To extend this to a multilevel linear model we add a random effect for each group $g = 1, \ldots, G$. Hence, for a set of *potentially* different predictors $\boldsymbol{z}_i$:

$$y_i = \boldsymbol{x}_i^T \boldsymbol{\beta} + \boldsymbol{z}_i^T \boldsymbol{\gamma}_g + \epsilon_i.$$

where the new parameters, $\boldsymbol{\gamma}$ take the role of group specific regression parameters and therefore are conditional on the group that y_i belongs to.

Model B is a special case of these models where only intercepts are fitted, sometimes referred to as a random effects ANOVA. Here,

$$y_i = \beta_0 + \gamma_g + \epsilon_i.$$

Since we actually fit this over multiple years worth of data, suppose y_{ij} is the i^{th} observation of the j^{th} year, then

$$y_{ij} = \beta_{0j} + \gamma_{gj} + \epsilon_{ij}.$$

We refer to β_{0j} as the population parameters as they represent the global value for year j and γgj are group-level parameters that denote the deviation from the population value for group g in year j.

Bibliography

British Gas. *How efficient is my boiler?* Website article, Accessed 27 March 2022, 2022. URL https://www.britishgas.co.uk/home-services/boilers-and-heating/guides/boiler-efficiency.html.

Paul-Christian Bürkner. Bayesian item response modeling in R with brms and Stan. *Journal of Statistical Software*, 100(5):1–54, 2021. doi: 10.18637/jss.v100.i05.

Paul-Christian Bürkner. brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software*, 80:1–28, 2017.

Bob Carpenter, Andrew Gelman, Matthew D Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. Stan: A probabilistic programming language. *Journal of statistical software*, 76(1), 2017.

Department for Business, Energy & Industrial Strategy. *Evidence gathering for electric heating options in off gas grid homes: Final Report*, 2019. URL https://www.gov.uk/government/publications/electric-and-bioenergy-heating-in-off-gas-grid-homes-evidence-gathering.

Department for Levelling Up, Housing and Communities. *English Housing Survey 2020 to 2021: headline report*, 2020a. URL https://www.gov.uk/government/statistics/english-housing-survey-2020-to-2021-headline-report.

Department for Levelling Up, Housing and Communities. *Local Authority Housing Statistics Data*. UK Government, Accessed 27 March 2022, 2020b. URL https://www.gov.uk/government/collections/local-authority-housing-data.

Department for Levelling Up, Housing and Communities. *Energy Performance Certificates for Buildings: Action Plan*, July 2018. URL https://www.gov.uk/government/consultations/energy-performance-certificates-in-buildings-call-for-evidence.

John Paul Gosling. SHELF: the Sheffield elicitation framework. In *Elicitation*, pages 61–93. Springer, 2018.

Stephen W Raudenbush and Anthony S Bryk. *Hierarchical Linear Models: Applications and Data Analysis Methods*, volume 1. Sage, 2002.

UK Government. *Energy Performance Certificate (EPC)*. UK Government, Accessed 4 April 2022, 2022. URL https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/5996/2116821.pdf.