ESCoE Economic Statistics Centre of Excellence

A COLLABORATION WITH Office for National Statistics

Energy Efficiency Improvements and Property Values: A Hedonic Analysis of Market Incentives in England and Wales

Robert Hill, Norbert Pfeifer and Miriam Steurer

ESCoE Discussion Paper 2023-12

July 2023

ISSN 2515-4664

DISCUSSION PAPER

Energy Efficiency Improvements and Property Values: A Hedonic Analysis of Market Incentives in England and Wales

Robert Hill, Norbert Pfeifer and Miriam Steurer

ESCoE Discussion Paper No. 2023-12
July 2023

Abstract

Increasing the energy efficiency of housing needs to be a key part of government strategies to reach Net Zero carbon emissions by 2050. In this paper we measure the market incentives of owners to improve the energy efficiency of residential properties and how these incentives differ by location and property type. By linking residential sales records for England and Wales with their Energy Performance Certificates (EPCs), we create a merged micro-level dataset providing transaction price, physical and locational characteristics, energy performance, recommended energy efficiency improvements, and associated costs at the level of individual properties. We also construct a proxy for plot size using the exact geographic location and distances to neighbouring properties. We then estimate a hedonic model to predict the property price increases if all EPC recommendations were implemented. Our results for the 2014-2022 period reveal significant differences in market incentives across regions and property types. On average, we find that 84.4% of the costs of EPC-recommended energy efficiency improvements are capitalised in property prices for flats, as compared with 59.4% for semis/terraces and 59.3% for detached houses, although significant differences exist between across regions. Subsidies targeted to regions and property types where market incentives are weakest could help reduce the cost of reaching Net Zero carbon emissions by 2050.

Keywords: Energy Performance Certificate; Energy Efficiency Improvements; Hedonic Model of Property Prices; Plot Size; Market Incentives

JEL classification: C43; Q52; Q54; R31

Robert Hill, University of Graz, Austria
Robert.hill@uni-graz.at

Published by:
Economic Statistics Centre of Excellence
King's College London
Strand
London
WC2R 2LS
United Kingdom
www.escoe.ac.uk

Energy Efficiency Improvements and Property Values: A Hedonic Analysis of Market Incentives in England and Wales

Robert Hill
University of Graz, Austria
robert.hill@uni-graz.at

Norbert Pfeifer
University of Graz, Austria
norbert.pfeifer@uni-graz.at

Miriam Steurer
University of Graz, Austria
miriam.steurer@uni-graz.at

July 3, 2023

Abstract

Increasing the energy efficiency of housing needs to be a key part of strategies to reach Net Zero carbon emissions by 2050. In this paper we measure the market incentives of owners to improve the energy efficiency of residential properties and how these incentives differ by location and property type. By linking sales records for England and Wales with their Energy Performance Certificates (EPCs), we create a merged micro-level dataset providing transaction price, physical and locational characteristics, energy performance, recommended energy efficiency improvements, and associated costs at the level of individual properties. We also construct a proxy for plot size using the exact geographic location and distances to neighbouring properties. We then estimate a hedonic model to predict the property price increases if all EPC recommendations were implemented. Our results reveal significant differences in market incentives across regions and property types. On average, we find that 84.4% of the costs of EPC-recommended energy efficiency improvements are capitalised in property prices for flats, as compared with 59.4% for semis/terraces and 59.3% for detached houses, although significant differences exist across regions. Subsidies targeted to regions and property types where market incentives are weakest could help reduce the cost of reaching Net Zero. (JEL. C43; Q52; Q54; R31)

Keywords: Energy Performance Certificate; Energy Efficiency Improvements; Hedonic Model; Housing Market; Plot Size; Market Incentives

We gratefully acknowledge funding from the UK's Economic Statistics Centre of Excellence (ESCoE) (Project 3.14). We thank Sanjiv Mahajan, Chris Jenkins, Tony Wilkins, Silja Guggisberg, and Sarah Sheppard for support throughout the project. We also thank participants at the ESCoE webinar (23 March 2023), the BANDAS seminar (29 March 2023) and the Wegener seminar (June 15, 2023) for helpful comments.

1 Introduction

> Almost 40% of the UK's energy consumption and carbon emissions come from the way our buildings are heated and used. Even comparatively small changes in energy performance and the way a building is used will have a significant effect in reducing energy consumption. (Department for Levelling Up, Housing and Communities, 2022)

In 2021, the UK government published its Net Zero Strategy which sets out a plan for the UK to reach net zero carbon emissions by 2050 (Department for Business, Energy and Industrial Strategy, 2022). In this paper we assess the market incentives for owners to increase the energy efficiency of their properties and how these incentives differ by location and property type. Such an analysis is an essential prerequisite to designing an effective strategy for increasing the energy efficiency of housing.

Since 2008, it has been mandatory for sellers of residential properties in the UK to provide a valid energy performance certificate (EPC) at the time of sale. EPCs are a useful mechanism to reduce the asymmetric information problem between buyers and sellers, increasing the likelihood that buyers will notice and appreciate measures that increase energy efficiency. This, in turn, should incentivise owners to invest in energy efficiency improvements as they are more likely to recoup the costs at a possible later sale.

A large literature has documented that properties with higher energy efficiency, other things equal, have higher prices (see section 2). The general approach to measuring the price impact of efficiency gains is to estimate a hedonic model of property prices where the energy efficiency label is included as a dummy variable. The consensus finding of this literature is that properties with higher energy efficiency labels are sold at a premium. However, the existing literature typically does not directly address owners' incentives to increase their properties' energy efficiency. A crucial input to any such discussion is the actual costs of these energy efficiency improvements. These costs can be compared with either the increase in the price of a property or the net present value of the energy cost savings. In equilibrium, these approaches should reach similar conclusions as the price of a property should rise by the net present value of the future cost savings.

Our focus is on measuring the price impact of energy efficiency improvements. To this end, we match the records for individual residential property transactions for England and Wales from the HM Land Registry with their EPCs. We then use a hedonic approach to estimate the impact of implementing the recommended energy efficiency improvements in the EPCs on property prices. We compute separate models for the three main property types (flats, semis/terraces, and detached houses) and the ten ITL-1 regions in England and Wales. The costs of the recommended energy efficiency improvements are provided directly in the EPCs.

We calculate the capitalisation rate of energy efficiency improvements (CREEI) at the level of individual properties. The CREEI is defined here as the percentage of the costs of energy efficiency improvements that are capitalised in the price of the property. For example, a CREEI of less than 100% implies the price of the property rises by less than the cost of the energy efficiency improvement, and hence that the improvement is not fully capitalised in property prices.

The HM Land Registry and EPC datasets provide a comprehensive list of characteristics of residential properties. These characteristics can be used in a hedonic model to minimise unobserved quality bias when estimating the impact of energy efficiency on property prices. In addition to the typical variables included in real estate models (i.e., price, size, age, and building type) and current and potential energy efficiency values, the combined dataset includes many additional property characteristics and features.

Plot size, however, is a key variable that is missing from both datasets.[1] We construct a proxy variable for plot size using an innovative approach based on the distances between neighbouring properties. These calculations are made using the exact geospatial locations of all addressable properties in England and Wales – a dataset of 40 million residential, commercial or government properties. Our plot size indicator variable is highly significant when included in our hedonic models.

We find that plot size can explain some apparent anomalies in the literature regarding the impact of energy efficiency on property values in the UK (Fuerst et al., 2015, 2016). When plot size is missing, the price effect of energy efficiency improvements can be substantially underestimated. This is especially true for detached houses, where the plot's share in the property's total value is much larger than for other property types.[2] Furthermore, the bias due to missing plot sizes will be greatest in areas where the proportion of land value (relative to the cost of the structure) is particularly high (e.g., around London or in the South East of England).[3]

The quality of the predictions from our hedonic models can also be affected by other omitted variables and measurement errors in the estimated efficiency levels of properties (Hyslop and Imbens, 2001). We perform two types of checks for bias in our hedonic results. The first is to

[1]The plot size is not usually recorded in English or Welsh title deeds. Property owners usually have access to a title plan that indicates the property's layout, but these plans do not specify exact boundaries nor who owns the hedge, wall, or fence between two properties UK Government (2023c). Also, Title plans are not available in a usable form for research purposes.

[2]Semis/terraces tend to have smaller and more uniformly sized plots. For flats, which usually do not have garden space, the internal square meter should be a sufficient description of the property size.

[3]This point ties in with the literature on the relative shares of land and building values in different markets (Bostic et al., 2007).

compare the price ratios of repeat-sales properties that underwent energy-efficiency improvements between sales with the predicted price ratios for these same properties obtained from our hedonic model. We find that the hedonic and repeat-sales results differ by only about 10%. Our second check focuses on controlling better for the structure's quality by using additional information in the EPCs. Again, our results suggest that any omitted variables biases are small.

We find significant differences in market incentives across regions and property types for the 2014-2022 period. On average, 84.4% of the costs of implementing the EPC-recommended energy efficiency improvements are capitalised in property prices for flats, as compared with 59.4% for semis/terraces and 59.3% for detached houses.[4]

Subsidies should be targeted to regions and property types where market incentives are weakest. The UK government's recently announced 12 billion pound "Help to Heat" scheme may be useful in this regard (Department for Energy Security and Net Zero, 2023). Our results can help policymakers design and target initiatives to maximise the reduction in carbon emissions per Pound of subsidy. Such careful use of government funds is a crucial element of any strategy for the UK to reach its Net Zero carbon emissions target by 2050.

The remainder of this paper is structured as follows. Section 2 is a literature review. Section 3 describes a general hedonic model of property prices. Section 4 develops the specific hedonic model we use to estimate the impact of energy efficiency improvements on property prices. The dataset is presented in section 5 and the results in section 6. Our robustness checks are in section 7 and the conclusion in section 8.

2 Relevant Literature

The literature exploring the impact of energy efficiency on property prices dates back to the 1980s. Early contributions include Johnson and Kaserman (1983), Laquatra (1986) and Dinan and Miranowski (1989). These authors found that higher energy efficiency does generate a price premium. However, the datasets used in these papers were small.

The field received a boost from the expansion of energy efficiency certification programs. These started first in the commercial real estate market. Eichholtz et al. (2010) considered the impact of Energy Star certification in the US office market. They found that Energy Star-certified properties commanded a 16% price premium. Similar studies of the commercial real estate market include Fuerst and McAllister (2011) and Kok et al. (2011).

[4] These are averages across the 10 ITL-1 regions shown in Table 10.

The first larger-scale empirical study considering the price impact of EPCs on residential real estate prices in Europe was Brounen and Kok (2011). Using a hedonic approach, they found that properties labelled A, B, or C commanded a price premium of 4% in the Dutch housing market. Since then, a number of papers have undertaken similar comparisons for other countries/regions, also finding that housing with higher energy efficiency labels command a price premium: see, for example Deng et al. (2012) for Singapore, Kahn and Kok (2014) for California, Hyland et al. (2013) for Ireland, Cerin et al. (2014) for Sweden, Fuerst et al. (2015) and Asproudis et al. (2023) for England, Fuerst et al. (2016) for Wales, Bruegge et al. (2016) for Florida, de Ayala et al. (2016) for Spain, Fuerst and Shimizu (2016) for Japan, Wahlström (2016) for Sweden, Walls et al. (2017) for the US, Kholodilin et al. (2017) for Berlin, Zhang et al. (2017) for China, Pride et al. (2018) for Alaska, Evangelista et al. (2020, 2022) for Portugal, Geske (2022) for England and Wales, and Taruttis and Weber (2022) for Germany. Finally, in a recent paper applying machine-learning methods to Zoopla list-price data, Asproudis et al. (2023) find that installing solar panels has a positive impact on the prices of detached houses in England.

In addition to the overall result that higher energy efficiency increases property prices, some other notable results emerge from these studies. For example, Taruttis and Weber (2022) find that in Germany, the price impact of improvements in energy efficiency is higher in rural areas. Fuerst et al. (2015) find that in England, the price premium is higher for flats and terraced dwellings than for detached and semi-detached properties. Aydin et al. (2020) find that improvements in energy efficiency are capitalised into house prices (i.e., the price premium exceeds or is at least equal to the net present value of energy cost savings). Geske (2022) finds that in England, the price premium exceeds the net present value of energy cost savings for high energy efficient properties while the premium is less than the energy cost savings for low energy efficient properties. Kholodilin et al. (2017) find that in Berlin, the price premium is about the same as the net present value of energy cost savings for owner-occupied properties, while the rent premium is much lower than the energy cost saving for rented properties. Cerin et al. (2014) find that in Sweden, the price premium is highest for properties built before 1960, while Kahn and Kok (2014) find that the price premium in California is higher in hotter regions where the cost of cooling is a larger share of total housing running costs. Bruegge et al. (2016) find that in Florida, the price premium is mainly for new builds, mostly disappearing for resales.

The impact of energy efficiency ratings may also depend on their market coverage. A potential selection issue exists if EPCs are voluntary. The impact of voluntary reporting of energy efficiency is explored by Fuerst and Warren-Myers (2018) for the rental market in Canberra, Australia. They find that improvements in energy efficiency would be valued more – especially at the cheaper end of the market – if reporting was mandatory. In the European Union, EPCs have been mandatory for transacted properties since 2009 (and have remained mandatory in

England and Wales after Brexit).[5]

Some authors caution that the price premiums estimated by the literature could be distorted by omitted variables (Aydin et al., 2020; Marmolejo-Duarte and Chen, 2022). The main consequential omitted variables here are plot size (for the UK market) and the quality of the housing structure/finishes. Repeat sales are sometimes used to circumvent or check the extent of the omitted variables problem (Fuerst et al., 2015, 2016; Aydin et al., 2020). However, unobserved quality may also have changed between the two sales since the energy improvement may be part of a broader renovation. We assess the extent to which these concerns over omitted variables are relevant to our results in section 7.[6]

3 Theoretical Model

The theoretical model presented here models Housing H as a differentiated good determined from the following elements:

$$H = H(Q^o, Q^u, E), \tag{1}$$

where Q^o refers to observed housing quality, such as the property's floor area, age and locational aspects. Q^u refers to the aspects of housing quality that remain hidden from the researcher, such as the plot size (in the UK), the quality of the building materials, kitchen, and other fittings. E denotes the energy efficiency.

Suppose now that households have the utility function

$$U = U(H, x), \tag{2}$$

where x denotes a composite consumption good. Properties can be either owner-occupied or rented. The cost of owner-occupying can be expressed as a user cost. The concept of user cost was developed by Hall and Jorgenson (1967) to provide an imputed rent for purchased capital used by firms in production. However, the concept can equally well be applied to the housing market to measure the cost of the services received by owner-occupiers (Poterba, 1984). In equilibrium, individuals should be indifferent between owning and renting such that the yearly

[5] However, this law is enforced more in some countries than in others. For example, de Ayala et al. (2016) note that EPCs are often not available for transacted properties in Spain.

[6] Aydin et al. (2020) propose an innovative approach to addressing the omitted variables problem using the oil price shock in 1974 to construct an instrument for a two-stage least squares regression. Such an approach can potentially control for some aspects of unobserved quality. Interestingly, in contrast to the Netherlands, the 1974 oil shock did not lead to the construction a more energy-efficient housing stock in England and Wales. Thus, the instrument approach used in Aydin et al. (2020) cannot be used with our dataset.

cost of renting (R) should equal the user cost of owner-occupying (uP).

$$R = uP \tag{3}$$

where P represents the property price and u is the per-dollar user cost calculated as follows:

$$u = r + \delta + \omega + \gamma - g, \tag{4}$$

The terms in (4) are defined as follows: r is the interest rate, δ is the rate of depreciation of the property, ω consists of repairs and maintenance, property taxes and average amortised transaction costs, γ is the risk premium for owning rather than renting, and g is the expected capital gain of owner occupying.[7]

The budget constraint of an individual considering investing in energy-efficient home improvements can be expressed as follows:

$$uP + x + C(\Delta E) - \frac{S(\Delta E)}{(1+r)^b} = M, \tag{5}$$

where ΔE denotes an improvement in energy efficiency, and $C(\Delta E)$ is the cost of this improvement. $S(\Delta E)$ is the per-period monetary saving from reduced energy usage, b is the number of periods for which the saving lasts, r is the interest rate, and M denotes income. The price of the composite consumption good (x) is normalised to 1.

Households choose H, x, and ΔE to maximise utility subject to their budget constraint. The first-order conditions yield the following solution for the equilibrium property price P:

$$P = P(Q^o, Q^u, E, r, \delta, \omega, \gamma, g, c, s, b, \bar{M}), \tag{6}$$

where c and s represent the parameters from the cost and savings functions in the budget constraint, and $\bar{M}$ denotes average income. We assume that average income is a sufficient statistic for modelling the impact of income on house prices.

Our objective is to calculate the change in property price associated with the change in the property's energy efficiency performance from E to E^+.

$$P(E^+) - P(E) = P(E + \Delta E) - P(E), \tag{7}$$

We then compare $P(E^+) - P(E)$ with $C(\Delta E)$ at the level of individual properties.

[7] A more general version of the user-cost formula is presented in Diewert (2009). Also, some versions of the user-cost method distinguish between the debt and equity components of housing wealth, where a mortgage interest rate is used for the debt part and a long-term government bond rate for the equity part.

We can distinguish between three types of variables. The first type concerns variables that we observe and can control for: Q^o, P, E and r. The second type consists of variables that we do not observe but can be assumed to be constant across properties at any given point in time: $\omega, \gamma, g, c, s, b, \bar{M}$. The fact that these variables are missing should not affect our estimates of ΔP much. The third type is the most problematic, as it refers to variables that are both missing from the dataset *and* can vary across properties. All (for the researcher) unobserved quality aspects of the property, Q^u, as well as the depreciation rate, δ, fall into this category.[8]

Consider two properties 1 and 2 that are identical in terms of all the observed characteristics (i.e., $Q_1^o = Q_2^o$), except that price and energy efficiency, are higher for the first property (i.e., $P_1 > P_2$ and $E_1 > E_2$). The problem is that it is likely that the unobserved qualities of property 1 are likely to also be larger than those for property 2 (i.e., $Q_1^u > Q_2^u$) (we ignore depreciation here). Thus, a hedonic approach will tend to overestimate the price increase resulting from an improvement in energy efficiency since it will attribute all the price differences to the higher energy efficiency when part of it is also due to higher property quality.

The better the dataset is able to describe the property (i.e., the more price-relevant variables are in Q^o rather than Q^u), the smaller this omitted variable bias should get. One aspect of this paper is to transform plot size, a crucial price-relevant variable, from a type Q^u into a type Q^o variable. Another aspect is to perform careful checks that help us estimate the robustness of our hedonic regression estimates (see section 7).

4 Hedonic Method

Hedonic regression can be used to estimate shadow prices on the characteristics of properties and for predicting the impact on a property of a change in its characteristics. The conceptual basis of the hedonic approach was developed by Lancaster (1966) and Rosen (1974). In a housing context, we estimate separately for each region and property type (flat, semi-terrace or detached house, or semi-terraced detached house) the following semi-log regression model on transacted properties as follows:

$$\ln(p_{t,n}) = \beta Q_n^o + \tau_t d_n + \theta E_n + \epsilon_{t,n}, \tag{8}$$

where $\ln(p_{t,n})$ is the log price of transaction n in period t, the vector Q_n^o describes the observed characteristics of property n including its location (the list of variables is discussed in section 5

[8] The depreciation of the structure is closely linked to the observed and unobserved quality of the property (Q^o and Q^u). The value share of land in the total value of the property also matters. Other things equal, a higher land share reduces the depreciation rate since land does not depreciate (Hill and Syed, 2016).

and section 6). d_n is a dummy variable that takes the value 1 when a property was transacted in year t, and zero otherwise. E_n is the property's energy efficiency level, and $\epsilon_{t,n}$ is an observation-specific error term with mean zero. For readability, the constant term is subsumed within the β vector. The parameters to be estimated are the characteristic shadow price vector β, the year-time-dummy vector τ, and the scalar θ, which measures the impact of energy efficiency on log price.

As a check for possible non-linearities in the impact of energy efficiency on price we also estimated a version of the hedonic model that interacts the energy efficiency parameter θ with the energy efficiency classes (A, B, C, etc.). As can be seen from the results in Appendix F, the estimated θ values are reasonably stable across all classes except class G, where the θ value is clearly lower. However, as shown in Table 2, class G accounts for only about 1% of the properties in our sample. Hence we use the simpler specification in (8).

Each EPC contains the property's current energy performance score, a list of recommended energy efficiency investments (including their estimated costs), and a potential energy score if all recommended energy improvements were realised. Once the hedonic model has been estimated, it can be used to predict the impact of energy efficiency improvements on property prices. Specifically, we use the estimated parameters from (8) to predict property prices. The first prediction is $\hat{p}_{t,n}$, the transaction price of property n given its current energy efficiency level (E_n) for each year t from 2014 to 2021. The second prediction is $\hat{p}^{+}_{t,n}$, the transaction price of property n if all recommendations in its EPC were implemented and its energy efficiency level was increased to (E^{+}).

These predicted prices are calculated as follows:

$$\hat{p}_{t,n} = \exp\left(\hat{\beta}Q^{o}_{n} + \hat{\tau}_{t}d_{n} + \hat{\theta}E_{n}\right), \tag{9}$$

$$\hat{p}^{+}_{t,n} = \exp\left(\hat{\beta}Q^{o}_{n} + \hat{\tau}_{t}d_{n} + \hat{\theta}E^{+}_{n}\right). \tag{10}$$

The expected price gain from implementing the recommended energy efficiency improvements listed in the property's EPC is given by $\hat{p}^{+}_{t,n} - \hat{p}_{t,n}$. Using the hedonic model to predict prices rather than focusing specifically on the estimated energy efficiency coefficient provides some protection against the possible effects of multicollinearity.

To evaluate the market incentive of property owners to undertake energy efficiency improvements the expected price gain needs to be compared with the cost of the improvement. Estimates of these costs, denoted here by $k_{t,n}$ are provided in the EPCs. The capitalisation rate of energy efficiency improvements (CREEI) is defined as follows:

$$CREEI_{t,n} = 100\left[\frac{\left(\hat{p}^{+}_{t,n} - \hat{p}_{t,n}\right)}{k_{t,n}}\right]. \tag{11}$$

A CREEI below 100% implies that energy efficiency improvements are less than fully capitalised in property prices. For example, if the CREEI is 90%, this means that when all recommendations in the EPC are implemented, the price of the property increases on average at a rate of 90 Pounds for every 100 Pounds of expenditure on energy efficiency improvements.[9]

When calculating the CREEI, the price of property n with the improvement, ($\hat{p}^{+}_{t,n}$), must be predicted since it is counterfactual. For a property actually transacted in period t, the price without the energy improvement, ($\hat{p}_{t,n}$), could be either the actual transaction price or the predicted price from the hedonic model. There are two reasons why it may be preferable to use the predicted price rather than the actual price. First, when calculating ($\hat{p}^{+}_{t,n} - \hat{p}_{t,n}$), there is no longer any need to restrict attention to properties that actually transacted in year t. We can consider all properties for which we have EPCs, independently of when (or even *if*) they were traded. Second, predicting both prices can partially correct for omitted variable bias. For example, suppose property n performs well in terms of the quality of the structure and finishes. This means that both $\hat{p}_{t,n}$ and $\hat{p}^{+}_{t,n}$ will tend to be too low. By taking the difference between these two predicted prices, these errors will partially offset each other. This approach of predicting both prices from the hedonic model is referred to as double prediction (or imputation) in the price index literature (Hill and Melser, 2008; Eurostat et al., 2013; Silver, 2018).

We derive aggregate measures of the CREEI by taking the median of the estimates in (11). We do this separately for each property type for each ITL-1 region.

5 Data

5.1 Matching EPC and PPD Datasets

HM Land Registry Transaction Price Dataset

The ”Price Paid Data” (PPD) contains all sales prices of properties in England and Wales that have been reported for registration following a change in property ownership (HM Land Registry, 2023). It contains each property’s sales price as stated on the transfer deed, its address, date of sale completion, property type (i.e., D = Detached, S = Semi-Detached, T = Terraced,

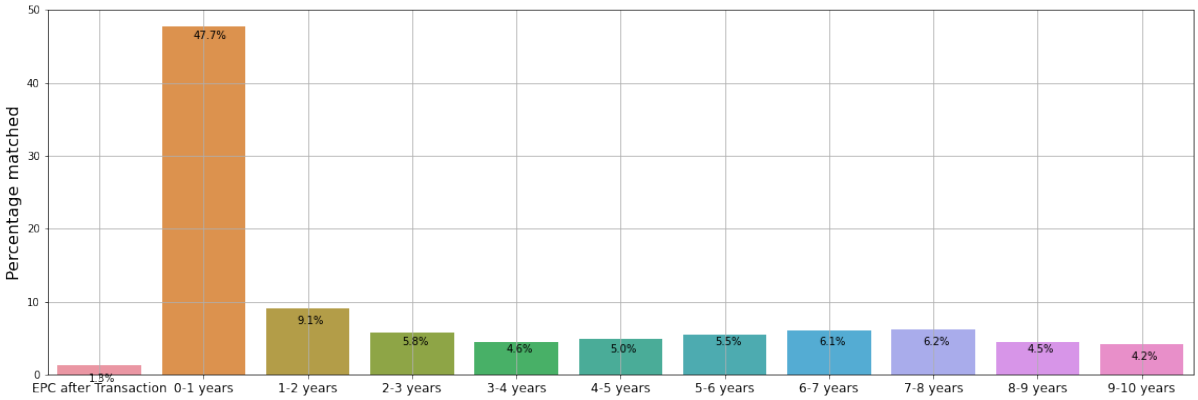
[9] We do not actually observe the increase in energy efficiency for each individual improvement. Some individual improvements will have higher capitalisation rates and others lower rates than the overall CREEI.

F = Flats/Maisonettes, O = Other), whether the property is a newbuild (”new” or ”old), and tenure type (i.e., L = leasehold, F = freehold). The PPD dataset also distinguishes between transaction types A and B, with Type A referring to ”Standard Price Paid” entries, which are single residential properties sold at market value. Type B refers to ”Additional Price Paid” records, which include transfers under the power of sale/repossession, lease-purchase, transfers to non-private individuals, and sales where the property type is classified as “Other” (UK Government, 2023b). As we are interested in residential properties sold at market value, we restrict the PPD dataset to transfer Type A properties.

Energy Performance Certificates (EPCs)

Energy performance certificates (EPCs) are based on individual assessments of properties' energy consumption and estimated energy losses (e.g., due to poor insulation). The assessments are performed by a trained assessor using an extensive checklist of property characteristics which are needed as inputs for the energy software/models. EPCs for buildings in England and Wales are published by (Department for Levelling Up, Housing and Communities, 2023). Apart from a few exceptions, valid EPCs have been required when selling domestic properties since September 2008 (Department for Levelling Up, Housing and Communities, 2023). Therefore, most EPCs are prepared in the course of a property sale, although EPCs can be up to ten years old at the time of the transaction (Office for National Statistics, 2021). Figure 1 illustrates the age of EPCs at the time of property transaction.

Figure 1: Age of EPC at the time of transaction



Note: This figure shows the matching pattern of PPD and EPC datasets for properties transacted in 2021. Overall, 88% of PPD properties could be matched with valid EPCs. A reason for the remaining 12% are commercial transactions, that are contained in the PPD, but not in the EPC dataset. Table E1 shows main statistics for the available characteristics between matched and un-matched. For about half of the transactions, the EPC certificate was issued in the year before the transaction.

Each EPC contains information about the energy efficiency class of the building during the energy assessment procedure, a list of proposed energy efficiency measures, and a potential energy-efficiency level if these recommendations were implemented. The EPC dataset contains a large set of property-specific information, for example, construction age, tenure type (rental-private, rental-social, owner-occupier, unknown), and built form (Flats, Terraces, Semis, Detached).

Matching of EPC and PPD Datasets

PPD and EPC records are linked via the exact addresses that are contained in the EPC and PPD records. Figure 2 illustrates that around 90% of yearly PPD transaction sales can be matched in this way. Failure to match can be caused by address discrepancies (when matching via addresses, we require exact address matches), commercial properties, or properties sold without EPCs. Although residential property owners are required (since 2008) to provide a valid EPC when selling their property, there are exceptions to this requirement for listed properties, vacant properties, properties smaller than 50 square meters, and properties with a valid building permit for demolition and rehabilitation (UK Government, 2023a).

EPCs are valid for ten years. Hence, some properties are sold with quite current EPCs, while others are sold with EPCs issued up to 10 years before the transaction. Figure 1 shows the age of the EPCs issued for the 661 650 properties sold in England and Wales in 2021. The longer the time interval between the EPC and transaction, the less confidence we can have in the estimated energy efficiency of a property, as it may have depreciated or been renovated in the meantime. To prevent outdated EPCs from influencing our results, we restrict the dataset to properties with EPCs issued in the two years prior to the transaction when estimating (8).[10] If properties are sold multiple times with the same EPC within this specified period, we keep only the first of those sales in our dataset since, presumably, the EPC best reflects the property's condition at the first sale after the EPC was issued.

[10]Changing this limit to one, five or ten years did not have much impact on the results.

Figure 2: Proportion of matched transactions

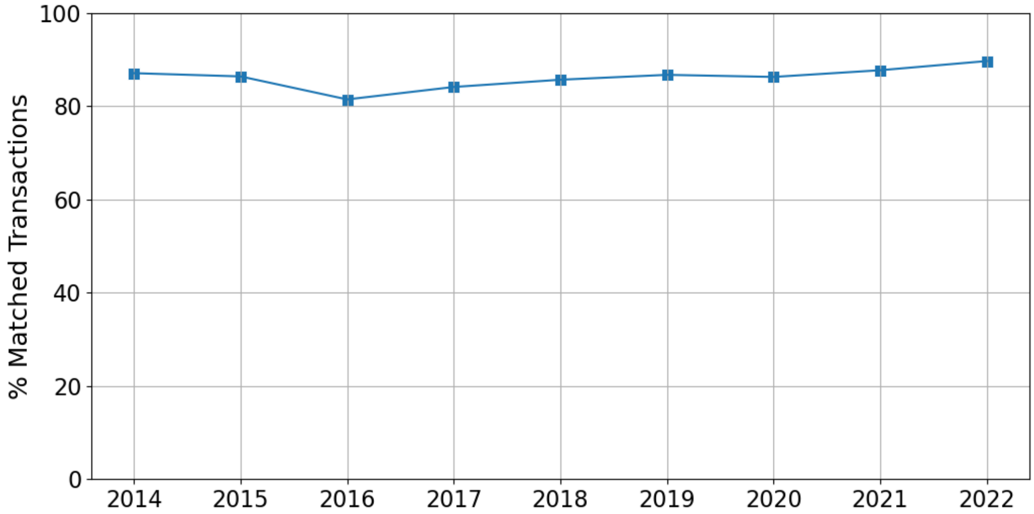


Figure 2 illustrates the percentage of transactions made for the years 2014 to 2022 that were matched to a corresponding EPC. Note that the transaction dataset includes commercial properties that are not included in the EPC sample.

5.2 Adding location-based data

PPD and EPC data contain addresses that allow us to classify properties regarding regional authority levels (i.e., county borders, districts, or postcodes). In our regressions, we use Local-Authority Districts (LAD) as locational fixed effects. To further describe the spatial aspect of the properties sold, we also integrate socioeconomic area information based on census data in the form of ”supergroups” and the “index of multiple deprivation”, both provided by the Office for National Statistics (ONS).

The eight classification labels of the supergroups are: “Affluent England”, “Business”, “Education and heritage centres”, “Countryside living”, “Ethnically diverse metropolitan living”, “London cosmopolitan”, “Services and industrial Legacy”, “Town and country living”, and “Urban settlements” (Office for National Statistics, 2023). Note that supergroup areas are defined independently of local authority levels Office for National Statistics (2023).

A smaller-scale socioeconomic classification based on census data is available via the ONS’s deprivation indices for England (Ministry of Housing, Communities and Local Government, 2019) and Wales (Welsh Government, 2019). The English indices, for example, rank all 32 844 small areas or neighbourhoods in England, called Lower-layer Super Output Areas. These deprivation indices provide a range of relative measures of deprivation for areas below the supergroup level (i.e., “lower-layer super output areas”). We use the “Index of Multiple Deprivation” which is based on seven different dimensions of deprivation: “Income Deprivation”, “Employment Deprivation”, “Education, Skills and Training Deprivation”, “Health Deprivation and Disability”, “Crime”, “Barriers to Housing and Services”, and “Living Environment Deprivation”.

Supergroups are included in our hedonic model as dummy variables. We include the multiple deprivation index as a numerical variable that, for ITL-1 regions in England, can take any value between 1 and 32 844.

5.3 Data cleaning and summary statistics

As linear regressions can be sensitive to outliers (Rousseeuw and Leroy, 1987), we use Cook’s distance to remove outliers in the dataset (Cook, 1979). To do so, we estimate hedonic models of log price separately for each property type (flat, semi-terrace, detached house) and ITL-1 region, taking into account the following characteristics: floor area, energy efficiency, duration (L/F), construction age band, Local-Authority District (LAD), built form, and plot-size indicator (for detached houses only). Table 1 provides the summary statistics for our dataset before and after removing observations with a Cook’s distance of greater than $4/(N - p - 1)$, where N stands for the number of observations and p for the number of parameters (Silver, 2022). This method classifies around 5% of the observations as outliers. In Appendix D, for comparison, we provide results computed without any data cleaning.

Table 1: Descriptive statistics of transaction data before and after cleaning

	Flat		Detached Houses		Semi/Terrace	
Cleaning	Before	After	Before	After	Before	After
Price (mean)	278,211	274,543	368,193	361,959	232,839	228,980
Floor Area (mean)	64.2	63.7	126.0	124.9	89.5	88.9
Construction age band (mode)	2007 onwards	2007 onwards	2007 onwards	2007 onwards	1930-1949	1930-1949
Built form detached-semi_detached(share)	40.3	40.4	97.9	98.2	53.1	53.6
Built form mid-terrace (share)	29.3	29.2	0.6	0.4	30.6	30.4
Built form end-terrace (share)	20.2	20.0	0.4	0.2	15.7	15.4
Built form missing (share)	10.2	10.5	1.1	1.2	0.6	0.6
Current Energy Efficiency (mean)	71.2	71.3	63.2	63.7	61.5	61.7
Potential Energy Efficiency (mean)	77.22	77.3	63.2	63.7	61.5	61.7
Region (share)						
North East	2.2		3.5		4.7	
North West	7.1		9.9		13.4	
Y. and the H.	4.1		8.6		10.4	
East Midlands	2.6		12.7		8.5	
West Midlands	5.7		10.0		9.9	
East	10.9		14.4		11.5	
London	35.7		2.5		10.0	
South East	20.1		19.6		16.2	
South West	9.7		13.5		10.4	
Wales	1.9		5.3		4.9	
Observations	766,538	730,579	1,592,490	1,510,864	3,553,395	3,397,605

Note: The columns indicated with “Before” show merged PPD-EPC records (with EPCs issued no more than 24 months before transaction). We applied Cook’s distance cleaning for each ITL-1 region separately, hence, the share of transactions from each region does nto change after cleaning.

The proportions of properties with each energy label are shown in Table 2. The energy efficiency is higher for flats, with level C being the mode. For semis/terraces and detached houses, the mode is level D, with 74% of semis/terraces and 64% of detached houses having efficiency levels of D or lower.

Table 2: Share of Energy Efficiency Level per Property Type

	Flats	Semi-Terraced	Detached
A	0.1	0.1	0.3
B	26.9	8.9	16.1
C	39.1	20.9	19.4
D	24.3	45.4	40.8
E	7.3	19.1	16.8
F	1.9	4.3	5.4
G	0.4	1.3	1.2

Note: Table 2 shows the share that belongs to each efficiency class using all market transactions between January 2008 and September 2022 that could be matched with a EPC within the last 2 years before transaction.

The impact of age on energy efficiency is shown in Figure 3. Three clear patterns can be discerned. First, newer buildings are more energy efficient. Second, flats are more energy efficient than semis/terraces which in turn are more energy efficient than detached houses. Third, this difference in the energy efficiency of the three property types falls over time. For properties constructed after 2007, on average there is no longer any difference in energy efficiency across property types.

Figure 3: Energy Efficiency By Age and Property Type

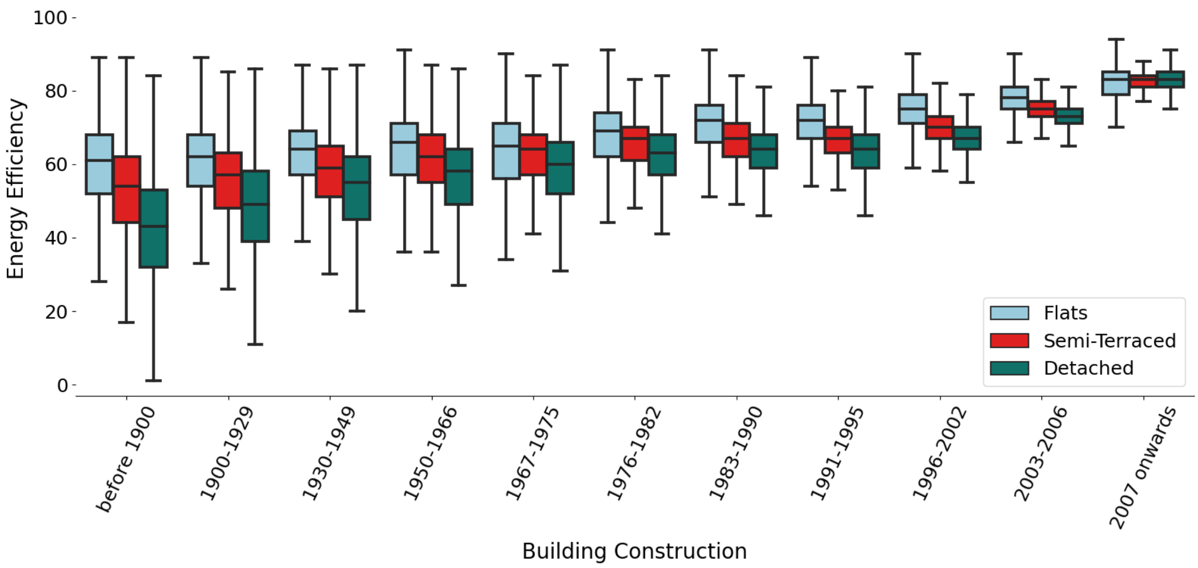


Figure 3 shows energy efficiency per property type for each building construction time range.

With regard to the government's Net Zero target for 2050, it is particularly important to identify those properties that have energy efficiency ratings below level C. Policies targeted at these properties could potentially provide the biggest gains in terms of reducing overall carbon emissions. From Table 3 it can be seen that the percentage of properties initially below level C is higher for semis/terraces and houses than for flats. Also, the percentages differ quite significantly across ITL-1 regions. What is especially striking is that the richest region, London, has the highest percentage of detached houses below level C (81.1%).

Table 3: Share of Properties for which the current energy efficiency is below level C

	North East	North West	Yorkshire and the Humber	East Midlands	West Midlands	East	London	South East	South West	Wales
Flats	51.8	38.5	41.1	39.4	48.1	45.3	50.1	46.2	47.1	40.1
Semi-Terrace	67.0	72.8	73.2	66.8	72.4	65.3	78.5	63.3	63.1	73.1
Detached	46.1	61.0	63.2	62.1	66.5	65.8	81.1	68.2	67.8	66.5

Note: Table 3 shows the percentage of properties of each type in each region that are below level C. These percentages are calculated as averages across the period 2014-2022.

5.4 Recommended energy efficiency improvements

The list of all energy improvements included in the EPCs is provided in Table 4, while the 10 most recommended energy efficiency improvements are shown in Table 5 for flats, in Table 6 for semi-terraces, and in Table 7 for detached houses. Low energy lighting for all fixed outlets is the most recommended improvement for flats and the third most recommended for semis/terraces and detached houses. This low-cost measure generates easy efficiency gains. The next most suggested improvement is 50 mm internal or external wall insulation. The most suggested improvement for semis/terraces and detached houses is the installation of solar photovoltaic panels followed by Solar water heating. Installing Solar panels is typically not an option for owners of flats who face different constraints than owners of semis/terraces and houses.

Table 4: List of Possible Recommendations for Improving Energy Efficiency in the EPCs

	Recommendation
1	50 mm internal or external wall insulation
2	Add additional 80 mm jacket to hot water cylinder
3	Cavity wall insulation
4	Change heating to gas condensing boiler
5	Change room heaters to condensing boiler
6	Condensing oil boiler with radiators
7	Draughtproof single-glazed windows
8	Fan assisted storage heaters
9	Fan assisted storage heaters and dual immersion cylinder
10	Flat roof insulation
11	Floor insulation
12	Flue gas heat recovery device in conjunction with boiler
13	Heat recovery system for mixer showers
14	High heat retention storage heaters
15	High heat retention storage heaters and dual immersion cylinder
16	High performance external doors
17	Hot water cylinder thermostat
18	Increase hot water cylinder insulation
19	Increase loft insulation to 270 mm
20	Increase hot water cylinder insulation
21	Install condensing boiler
22	Insulate hot water cylinder with 80 mm jacket
23	Low energy lighting for all fixed outlets
24	Party wall insulation
25	Replace boiler with new condensing boiler
26	Replace heating unit with condensing unit
27	Replace heating unit with mains gas condensing unit
28	Replace single glazed windows with low-E double glazing
29	Replacement glazing units
30	Replacement warm air unit
31	Room-in-roof insulation
32	Secondary glazing to single glazed windows
33	Solar photovoltaic panels, 2.5 kWp
34	Solar water heating
35	Solid floor insulation
36	Suspended floor insulation
37	Time and temperature zone control
38	Upgrade heating controls
39	Wind turbine
40	Wood pellet stove with boiler and radiators

Table 5: 10 Most Frequent Recommendations: Property Type – Flat

Recommendation	Share	Mean Costs (at 2011 prices)
Low energy lighting for all fixed outlets	59.50	25
50 mm internal or external wall insulation	37.99	9 000
Cavity wall insulation	33.81	1 000
Replace boiler with new condensing boiler	20.50	2 600
Solid floor insulation	18.09	5 000
Replace single glazed windows with low-E double glazing	16.53	4 900
High heat retention storage heaters and dual immersion cylinder	15.63	1 500
Increase loft insulation to 270 mm	14.73	225
High heat retention storage heaters	14.29	1 500
Upgrade heating controls	11.09	400

Note: Here "share" denotes the percentage of EPCs across our sample period (2014-2022) that include this recommendation. For example, 59.5% of EPCs for flats suggest the installation of low-energy lighting. The column "mean cost" estimates the average cost across the whole sample period (2014-2022) as stated in the individual EPCs (i.e., costs have not been uprated – see section 5.6).

Table 6: 10 Most Frequent Recommendations: Property Type – Semis/Terraces

Recommendation	Share	Mean Costs (at 2011 prices)
Solar photovoltaic panels, 2.5 kWp	96.72	5 000
Solar water heating	91.06	5 000
Low energy lighting for all fixed outlets	57.59	25
50 mm internal or external wall insulation	33.67	9 000
Solid floor insulation	30.84	5 000
Replace boiler with new condensing boiler	27.71	2 600
Suspended floor insulation	27.70	1 000
Cavity wall insulation	22.86	1 000
Upgrade heating controls	18.44	400
Increase loft insulation to 270 mm	17.02	225

Note: Here "share" denotes the percentage of EPCs across our sample period (2014-2022) that include this recommendation. For example, 96.7% of EPCs suggest the installation of solar photovoltaic panels, 2.5 kWp. The column "mean cost" estimates the average cost across the whole sample period (2014-2022) as stated in the individual EPCs (i.e., the costs have not been uprated – see section 5.6).

Table 7: 10 Most Frequent Recommendations: Property Type – Detached House

Recommendation	Share	Mean Costs (at 2011 prices)
Solar photovoltaic panels, 2.5 kWp	95.48	5 000
Solar water heating	83.11	5 000
Low energy lighting for all fixed outlets	53.82	25
Solid floor insulation	37.11	5 000
Replace boiler with new condensing boiler	29.06	2 600
Suspended floor insulation	21.96	1 000
Cavity wall insulation	19.46	1 000
Increase loft insulation to 270 mm	17.01	225
Floor insulation	13.38	1 000
Upgrading heating controls	12.62	400

Note: Here "share" denotes the percentage of EPCs across our sample period (2014-2022) that includes this particular recommendation. I.e., for detached houses 95.5% of EPCs recommend the installation of solar photovoltaic panels, 2.5 kWp. The column "mean cost" estimates the average cost across the whole sample period (2014-2022) as stated in the individual EPCs (i.e., the costs have not been uprated – see section 5.6).

5.5 Including an estimate of plot size

Plot size is an essential determinant of the price of single-family homes. However, neither the PPD- nor the EPC dataset includes any information about the size of the land the property is standing on. The lack of this variable can lead to significant missing characteristics bias in the hedonic regression models, especially for detached houses. To address this shortcoming, we construct our own proxy measure for plot size based on the distance of a property from other buildings.

The Unique Property Reference Number (UPRN) dataset provides us with the data we need to determine the precise geographic location (in terms of latitude and longitude) of every structure in England and Wales - independently of whether they have been sold, or whether they are categorised as residential, commercial, industrial, or government building. Forty million properties are covered in this way.[11]

Our proxy measure is based on the realisation that plot size is positively related to the distance between properties, as large plots are only possible if neighbouring buildings are distant. Thus, the greater the distance to neighbouring properties, the larger a property's (potential) plot. Here, we construct our indicator variable for plot size as the mean distance of a property address to its ten closest neighbours. However, our hedonic regression results are quite robust to the exact number of UPRN neighbours used (as long as some distance or distance-squared variable is

[11] The exact size of the dataset is 39,919,499, and it can be accessed at https://uprn.uk/.

included). Figure 4 graphically illustrates the concept behind our proxy plot variable.

Figure 4: Calculation of plot size approximation

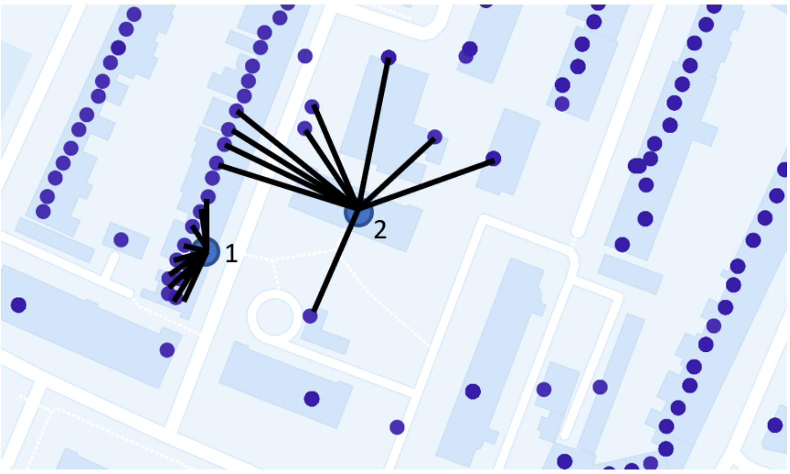


Figure 4 illustrates the calculation of the plot size proxy for detached houses. Each unique address is marked as a unique reference point (the small blue circles). Black solid lines depict the distances to the ten closest neighbours. Reference point 1 (left large blue circle) has a lower average distance to its ten nearest neighbours than reference point 2 (right large blue circle). Property 1's potential plot size is thus larger than that of property 2.

Figure 5: Plot-Indicator Measurements for Detached Houses

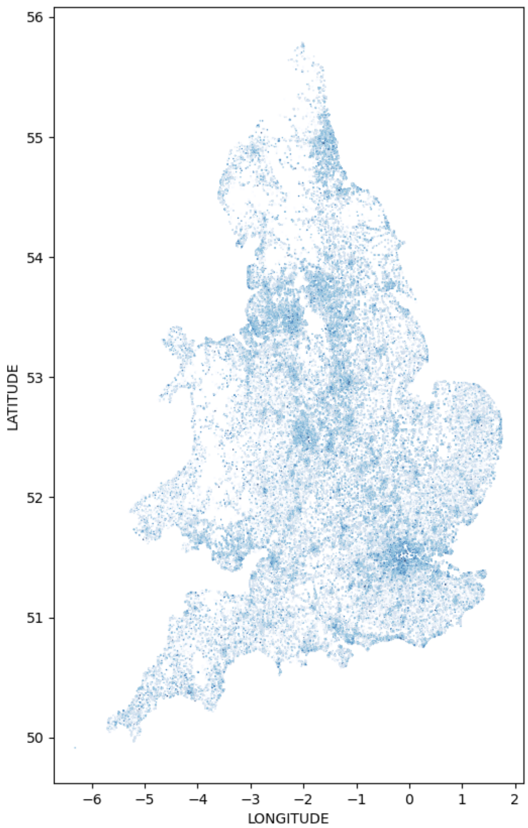
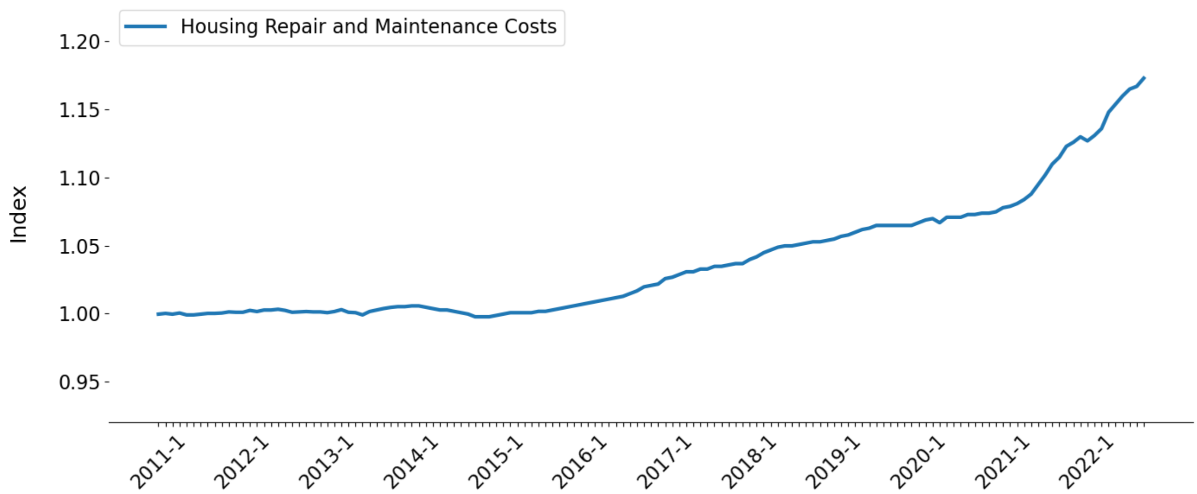


Figure 5 illustrates the plot-size indicator variable for detached houses in England and Wales. Darker dots indicate closer distance to neighbours.

5.6 Uprating of costs

Starting in 2011, EPCs in England and Wales include individual energy improvement recommendations and estimated costs. However, these costs have not increased since 2011, which is clearly implausible. Hence we uprate the costs stated in the EPCs for years after 2011 using an index constructed by splicing together two ONS housing repairs and maintenance cost indices. We inflate EPC costs between 2011 and 2014 using the RPI: housing: repairs and maintenance charges index.[12] From 2014 onwards we use the construction price index: Housing repairs and maintenance.[13] Figure 6 represents this spliced cost index. As shown in Figure 6, the cost of housing repairs and maintenance increased by 25% between 2011 and 2022. By comparison, the official House Price Index produced by ONS rose by 44% over this same period. Without considering these cost increases, we would have significantly overstated the market incentives for energy efficiency improvements in the later years of our sample. We provide CREEI estimates based on the original EPC-stated costs in Appendix B.

Figure 6: Changes in the Cost of Housing Repair and Maintenance (2011-2022)



Note: The cost index splices together two ONS indices. Prior to 2014 the RPI: housing: repairs and maintenance charges index is used. From 2014 the construction price index: Housing repairs and maintenance is used.

[12] https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/dobt/mm23

[13] https://www.ons.gov.uk/businessindustryandtrade/constructionindustry/datasets/interimconstructionoutputpriceindices

6 Results

6.1 Results from hedonic model

The hedonic model in (8) is run separately for each ITL-1 region with data from 2008-2022.[14] The dependent variable is the logarithm of the transaction price. Apart from energy efficiency scores, we include the following characteristics in the hedonic regression: floor area (in square meters), deprivation (as a rank index on relative deprivation), supergroup levels,[15] property type sub-levels, construction age bands,[16] tenure levels (i.e., owner-occupied or private rental or social rental), duration (leasehold or freehold), built form (detached, semi-detached, end-terrace, mid-terrace) and year of sale.

Most of the literature measures the impact of energy efficiency on house prices by including dummy variables in the hedonic model for different energy classes (for the case of EPCs in Europe, these classes are A, B, C, D, E, F and G). In papers that include the actual energy efficiency score from the EPC, there is some debate in the literature on whether energy efficiency should enter the hedonic regression model in linear or log form. For example, Cerin et al. (2014) and Aydin et al. (2020) include the log of energy efficiency, while Wahlström (2016) includes the log of energy consumption per square meter (which is an inverse function of energy efficiency). By contrast, Kholodilin et al. (2017) and Taruttis and Weber (2022) include energy consumption per square meter linearly in the hedonic model. We use the linear specification for two reasons. First, our results in Appendix F show that the θ parameter in our hedonic model is reasonably stable when we interact it with the energy efficiency class. Second, a linear specification provides a better fit of the data based on an examination of the scatter plots of log price on energy efficiency as compared with log price on log energy efficiency.

The regression results are illustrated in Table I1 for flats, in Table I2 for semis/terraces, in Table 8 for detached houses without the plot size proxy, and in Table 9 for detached houses with the plot size proxy. Each Table shows the estimated parameters from the hedonic models of each of the 10 ITL-1 regions. The adjusted R-squared coefficients of the hedonic models range from 0.66 to 0.88, showing that the characteristics included in the hedonic models explain a substantial part of the variation in property prices. All control variables have the expected signs and are highly significant.

[14] We present results from 2014 to September 2022. The earlier years are needed to obtain a large enough repeat-sales sample for the robustness checks undertaken in section 7.

[15] A classification based on census results into regional types, differentiating between "Rural", "Cosmopolitans", "Ethnicity", "Multicultural", "Urbanites", "Suburbanites", "Constrained", and "Hard-pressed"

[16] This variable describes the year of construction as follows: before 1900, 1900-1929, 1930-1949, 1950-1966, 1967-1975, 1976-1982, 1983-1990, 1991-1995, 1996-2002, 2003-2006, 2007 onward

The coefficients on energy efficiency are positive and significant at the 0.1% significance level for flats and semis/terraces and for all regions except London for detached houses without plot size proxy (where the shadow price on energy efficiency is negative). This counter-intuitive result is similar to what is found by Fuerst et al. (2015). It strongly indicates a missing variable bias that arises because plot size is not included as a control variable in the hedonic model. Plot size is particularly important for detached houses in London where the value of land relative to building structure is especially high. Thus, it is crucial that a proxy for plot size is included for detached houses in the hedonic model, particularly in areas where land values are high. From Table 9, it can be seen that when the plot size proxy is included, the shadow price on energy efficiency is positive and significant at the 0.1% level in London. Also, the leasehold coefficient switches from positive to negative, in line with what one would expect. In addition, the inclusion of the plot size proxy increases the adjusted R^2, indicating that the hedonic model is better able to explain price variations with the plot variable included in the hedonic regression (see the comparison between Table 8 and Table 9).

Comparing the shadow prices across ITL-1 regions we find that the North East has the highest shadow price of energy efficiency for flats and detached houses, while for semis/terraces, Wales has the highest shadow price. For all three property types, London has the lowest shadow price of energy efficiency. Overall, increases in energy efficiency seem to have a larger impact on property prices in the poorer ITL-1 regions. One possible interpretation of this finding is that in more expensive regions property prices depend less on the physical characteristics of properties and more on the location (and plot size in the case of detached properties).

Figure 7a to Figure 7c illustrate how well the hedonic models are able to predict actual transaction prices. For each transacted property, the x-axis shows the actual price, and the y-axis shows the predicted price from the hedonic model. The observations are all reasonably closely clustered around the 45-degree line for all three property types. The mean absolute percentage errors (MAPE) are 17.8%, 15.3% and 13.9% for flats, semi-terraces and detached houses, respectively.

Table 8: Regression Output for Detached Houses (1) – no plot variable

	North East	North West	Yorkshire and the Humber	East Midlands	West Midlands	East	London	South East	South West	Wales
Constant	11.1480***	11.3327***	11.1983***	11.4174***	11.6153 ***	11.5286***	12.0088***	11.8918***	11.8741***	11.3837
	(0.018)	(0.013)	(0.015)	(0.009)	(0.012)	(0.011)	(0.053)	(0.011)	(0.011)	(0.016)
Energy Efficiency	0.0031***	0.0014***	0.0013***	0.0019***	0.0012***	0.0012***	-0.0003**	0.0002***	0.0008***	0.0017***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Floor Area	0.0064***	0.0062***	0.0060***	0.0060***	0.0059***	0.0055***	0.0053***	0.0055***	0.0058***	0.0055***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Deprivation	9.64e-06***	9.329e-06***	1.002e-05***	6.649e-06***	6.983e-06***	5.904e-06***	1.029e-05***	5.963e-06***	6.577e-06***	0.0002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
House	Base	Base	Base	Base	Base	Base	Base	Base	Base	Base
Bungalow	0.0862***	0.0317***	0.0213***	0.0361***	0.0405***	0.0071***	-0.0797***	-0.0174***	0.0274***	0.0038*
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)	(0.002)
Freehold	Base	Base	Base	Base	Base	Base	Base	Base	Base	Base
Leasehold	-0.0352***	-0.0039**	-0.0119***	-0.0350***	-0.0310***	-0.0649***	0.0031	-0.1216***	-0.0580***	-0.0065
	(0.003)	(0.001)	(0.003)	(0.004)	(0.003)	(0.006)	(0.020)	(0.007)	(0.007)	(0.005)
Supergroup	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Construction Age	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tenure	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Built Form	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Local Authority	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	52,993	148,989	129,822	191,754	151,503	217,963	37,112	296,388	205,036	79,304
Adjusted R^2	0.78	0.76	0.80	0.82	0.82	0.86	0.84	0.83	0.75	0.74

The table summarizes the output for the hedonic models for detached houses. The dependent variable is the logarithm of the transaction price. The model is run separately for each ITL-1 region with data from 2008-2022. All included control variables have the expected sign and are highly significant. Note: Asterisks representing significance. * p-value <= 0.05; ** p-value <= 0.01, *** p-value <= 0.001.

Table 9: Regression for Detached Houses (2) – including plot variable

	North East	North West	Yorkshire and the Humber	East Midlands	West Midlands	East	London	South East	South West	Wales
Constant	10.9712***	11.1393***	11.0452***	11.2996***	11.4265***	11.3716***	11.7228***	11.6824***	11.7292***	11.2490
	(0.018)	(0.013)	(0.015)	(0.009)	(0.011)	(0.010)	(0.010)	(0.000)	(0.010)	(0.016)
Energy Efficiency	0.0039***	0.0024***	0.0021***	0.0025***	0.0021***	0.0018***	0.0003***	0.0012***	0.0015***	0.0023***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Plot Indicator	3.6729***	3.6114***	3.9902***	2.9775***	2.8186***	2.9851***	9.8124***	4.5580***	2.5873***	1.2735***
	(0.065)	(0.035)	(0.042)	(0.028)	(0.024)	(0.023)	(0.154)	(0.023)	(0.021)	(0.019)
Floor Area	0.0060***	0.0059***	0.0056***	0.0057***	0.0056***	0.0051***	0.0046***	0.0049***	0.0055***	0.0054***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Deprivation	9.643e-06***	9.43e-06***	1.022e-05***	6.849e-06***	6.983e-06***	6.324e-06***	9.941e-06***	6.387e-06***	6.884e-06***	0.0002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
House	Base	Base	Base	Base	Base	Base	Base	Base	Base	Base
Bungalow	0.0689***	0.0186***	0.0058***	0.0245***	0.0295***	-0.0053***	-0.1048***	-0.0347***	0.0170***	0.0022
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)	(0.005)
Freehold	Base	Base	Base	Base	Base	Base	Base	Base	Base	Base
Leasehold	-0.0342***	-0.0019	-0.0115***	-0.0304***	-0.0291***	-0.0656***	-0.0193	-0.1229***	-0.0575***	-0.0063
	(0.003)	(0.001)	(0.003)	(0.004)	(0.003)	(0.006)	(0.019)	(0.007)	(0.007)	(0.000)
Supergroup	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Construction Age	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tenure	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Built Form	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Local Authority	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	51,986	146,070	127,668	189,747	149,325	215,444	36,279	292,949	201,972	77,050
Adjusted R^2	0.79	0.79	0.81	0.83	0.83	0.87	0.85	0.85	0.76	0.75

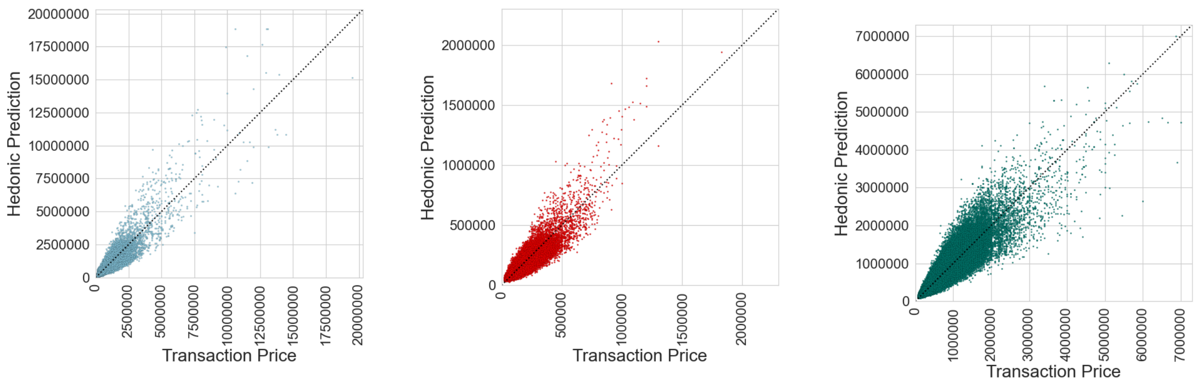
The table summarizes the output for the hedonic models for detached houses when we include the plot indicator variable for each property in addition to the variables described in Table I1. Again, the model is run separately for each ITL-1 region and the dependent variable is the logarithm of the transaction price. As we have not yet estimated the plot indicator variable for all years, so far only data for the years (2017 to 2021) have been taken into account. All control variables have the expected sign and are highly significant. Note that, compared to the regression output in Table 8, the estimated energy efficiency coefficient increased for all regions due to the inclusion of the plot indicator variable.
The table also shows that the plot-size variable is not only highly significant (at least for the detached house models), but that its presence can also have a large impact on the estimated energy efficiency parameter value. This is particularly the case in the London region, where the value share of land (relative to structure) is highest. Note: Asterisks representing significance. * p-value <= 0.05; ** p-value <= 0.01, *** p-value <= 0.001.

Figure 7: Prediction Performance of Hedonic Model (2014-2022)

(a) Flats

(b) Semi-Terraced Houses

(c) Detached Houses with the plot size proxy included



Note: The x-axis shows the actual transaction price of a property, and the y-axis its predicted price from the hedonic model. The dashed line in each panel is a 45-degree line. A dot is located above the line when its predicted price is higher than the transaction price and below the line when it is below the transaction price.

6.2 Does it pay to implement the EPC recommendations?

We can now investigate whether it pays to invest in energy-saving home improvements. Specifically, we estimate the average Capitalisation Rate of Energy Efficiency Improvements (CREEI), as defined in (11), separately for each property type (i.e., flats, semis/terraces, and detached houses) for each of the ten ITL-1 regions.

EPC recommendations are provided for each property based on its property type, installed energy-saving features, and current energy efficiency level. Figures C1 to C3 in Appendix C are pie charts that show the current energy efficiency labels (A to G) for each ITL-1 region for each property type in 2019. The pie charts for flats and semis/terraces both look quite similar across regions. By contrast, the differences across regions are much larger for detached houses. In particular, 65% of detached houses in the North East are at level C or higher, compared with only 42% in London.

The average CREEI estimates over the period 2014-2022 for each ITL-1 region and property type are shown in Table 10. Also provided are bootstrapped 95% confidence intervals for the average CREEIs.[17] For flats, most of the CREEIs are below 100% implying that energy efficiency improvements are less than fully capitalised in property prices. The East Midlands is the only region with an average CREEI above 100%. Its average CREEI is 103%. This means that the prices of flats in the East Midlands increase on average at a rate of 103 Pounds for every 100 Pounds of expenditure on energy efficiency improvements, when all recommendations in the EPC are implemented. The average CREEIs for flats for the other regions lie in the range 72%

[17] Graphs of the bootstrapped results are provided in Appendix G.

to 96%. Hence, for all but one region, energy efficiency improvements are not fully capitalised in property prices.[18]

The CREEIs for semis/terraces in Table 10 are clearly lower than for flats. Hence, the financial incentives for owners of semis/terraces to implement all the recommendations in EPCs are lower than for flat owners. This finding is probably at least partly a consequence of the more substantial recommendations typically suggested for semis/terraces compared to flats. For example, most flat owners would not be able to install solar panels or increase loft insulation, while this is generally possible for semis/terraces. The average cost of implementing all EPC recommendations for flats in 2022 is £4 874; for semis/terraces it is £15 733; and for detached houses the average cost of implementing all recommendations is £14 316. For semis/terraces, the highest CREEI of 65% is observed in the Northwest region.

The most striking result in Table 10 for detached houses without the plot size proxy is the negative CREEI for London. This would mean that increasing energy efficiency lowers the price of a property. Our interpretation is that this is an erroneous result caused by plot size in London being positively correlated with price but negatively correlated with energy efficiency. Including the plot size proxy in the hedonic model increases all the average CREEIs. Furthermore, the CREEI is now positive for London, although it is still much lower than in the other regions (24%). The large premium on location and plot size in London seems to dwarf the benefits of energy improvements.[19] The CREEIs of other regions lie in the range 55% and 88%.[20]

[18]We ignore here the savings in terms of lower energy costs. However, in equilibrium, the increase in property price should equal the net present value of the expected reduction in future energy costs. An additional cost of implementing energy efficiency improvements that we do not model is the associated inconvenience of undertaking any kind of renovation.

[19]This finding may also be linked to the much lower average energy efficiency of detached houses in London, as shown in Figure C3.

[20]A caveat to our results is that we estimated the CREEIs on the basis that all recommendations suggested in individual EPCs are implemented. In practice, some of these recommendations will be more cost-effective than others.

Table 10: Capitalisation Rate of Energy Efficiency Improvements (CREEI)

	North East	North West	Yorkshire and the Humber	East Midlands	West Midlands	East	London	South East	South West	Wales
Flats	$\boxed{87.8}$	96.0	80.6	103.2	80.0	72.3	91.1	73.7	76.0	83.4
Lower Quantile (2.5%)	78.3	87.0	69.5	90.3	70.7	65.6	83.9	67.8	68.8	69.4
Upper Quantile (97.5%)	98.3	105.6	93.6	117.2	88.8	79.7	98.2	79.7	83.6	97.7
Semi-Terraced Properties	47.0	65.0	63.2	61.9	60.2	56.9	64.7	60.4	51.1	63.2
Lower Quantile (2.5%)	45.2	64.0	61.9	60.5	58.9	55.4	60.8	58.7	49.3	61.8
Upper Quantile (97.5%)	48.8	66.2	64.4	63.4	61.6	58.6	67.8	62.0	52.9	64.9
Detached Houses	70.8	38.4	38.8	50.3	36.0	42.8	-19.6	8.9	31.6	41.2
Lower Quantile (2.5%)	65.9	34.8	35.1	47.7	32.4	39.3	-36.9	5.3	28.0	36.9
Upper Quantile (97.5%)	76.1	41.8	42.7	52.7	39.4	46.6	-2.9	12.5	34.7	45.1
Detached Houses with plot size	88.1	64.8	59.8	64.7	61.4	63.9	24.0	56.4	54.8	55.0
Lower Quantile (2.5%)	83.3	61.1	56.3	62.0	57.9	60.6	9.1	52.8	51.6	51.0
Upper Quantile (97.5%)	93.4	68.2	63.2	67.3	64.5	67.0	39.9	60.2	58.3	58.6

Note: The formula for computing the CREEI is provided in (11). The average CREEI for the North East is $\boxed{87.8\%}$. This implies that when all recommendations in the EPC are implemented, the price of a flat in the North East rises on average at a rate of 87.8 Pounds for every 100 Pounds of expenditure on energy efficiency improvements.

In Appendix H we extend the analysis to consider how the results change if we restrict the sample of properties to those that are initially below level C. Generally, we find that the CREEIs are higher for properties that start below level C (for semi-terraces and detached houses) than for properties that start at or above level C. Even then, the CREEIs are typically still below 100.

7 Robustness Checks

7.1 Repeat-sales of efficiency-improved properties as an omitted variables robustness check

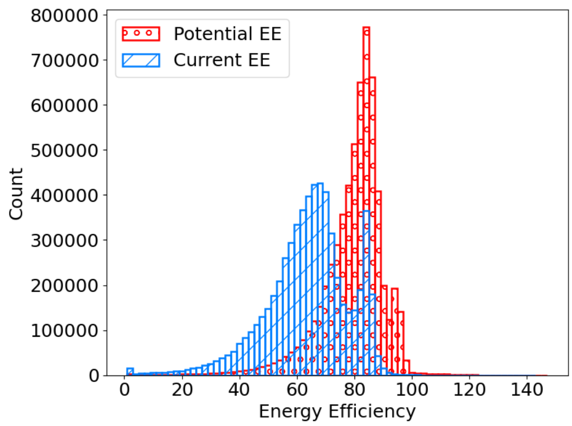
Repeat-sales samples can provide a robustness check on our hedonic results. Let $p_{t,n}(E_{t,n})$ and $p_{t+k,n}(E_{t+k,n})$ denote the actual transaction prices of property n in years t and $t+k$. We focus on repeat sales for which a new EPC is issued between the two sales. Hence in year $t+k$, the property has a different level of energy efficiency than in year t. 403 413 properties (out of 7 046 478) satisfy this requirement.[21] We further limit our repeat-sales sample to those properties for which property size and construction year stayed the same (so we can rule out property extensions and re-builds as the reason for the price change). Additionally, we exclude properties for which the

[21] We initially started with 10 117 689 sales, which account for all matched EPCs in our sample. After deleting all non-market transactions and property types that fall not within the following categories: flats, semis/terraces, and detached houses, and accounting for sales where the inspection date of the EPC was after the transaction date, we end up with 7 046 478 single transactions.

two transactions occur within 12 months to exclude properties with a high likelihood of being purchased, renovated, and resold by a *builder.[22] The repeat-sales sample we ended up with contains 8 753 properties sold twice between 2008 and 2022. Figure 8b depicts the distribution of energy efficiency at the first and the second sale from the repeat sales sample. It can be seen that the first sales generally have lower energy efficiency levels than the second sales.

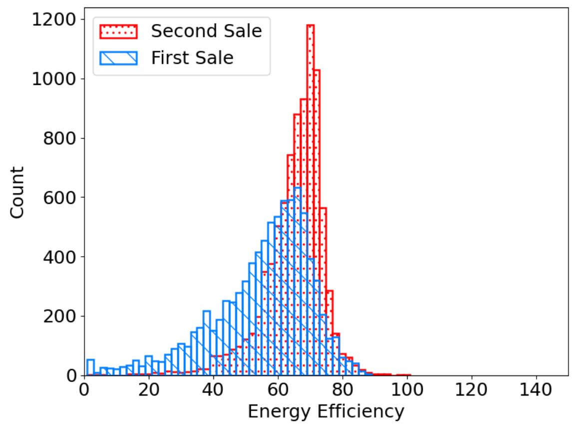
Figure 8: Sample differences in energy efficiency

(a) Current and potential energy efficiency – full sample



The median current- and potential energy efficiency for the whole sample is 65 and 81 respectively.

(b) Energy efficiency of first and second sales - repeat sales sample



The median current energy efficiency value of the first- and second sales is 58 and 66 respectively.

We can now compare the actual price change $p_{t+k,n}(E_{t+k})/p_{t,n}(E_{t,k})$ with the predicted price change from the hedonic models for periods t and $t+k$ for all properties in the two-EPC-repeat-sales sample. We make this comparison in two ways. The average difference as a ratio is measured by K:

$$K = \prod_{n=1}^{N} \left(\frac{\hat{p}_{t+k,n}(E_{t+k,n})}{\hat{p}_{t,n}(E_{t,n})} \bigg/ \frac{p_{t+k,n}(E_{t+k,n})}{p_{t,n}(E_{t,n})} \right)^{1/N}. \tag{12}$$

One must be careful interpreting the K coefficients as the hedonic and repeat-sales results are subject to different biases. Three types of bias may affect the hedonic estimates of the impact of energy efficiency improvements on property prices. First, the unobserved quality aspects of the structure may be positively correlated with energy efficiency. This may cause the hedonic model to overestimate the impact of energy efficiency improvements on property prices. The repeat-sales price changes may also face this problem because when energy improvements are undertaken, some unobserved quality aspects (such as installing a new kitchen or bathroom) may be simultaneously improved. The literature has tended to assume that quality bias affects

[22] Omitting higher frequency repeat sales is a standard approach when constructing repeat-sales price indices. For example, the Case-Shiller index excludes repeat sales that occur within six months (S&P Dow Jones Indices, 2022) and Aydin et al. (2020) exclude repeat sales within 12 months. If we allow a minimum of six months between two sales for one property, then the repeat sales sample increase to 9403 (+650) and the overall K index (see below) reduces from 0.90 to 0.88.

repeat-sales price changes less than those obtained from hedonic models. However, this is not necessarily the case.

The second and third biases apply only to the hedonic result. The plot size is missing in the English and Welsh data. To the extent that plot size correlates with energy efficiency, this can again bias the results obtained from hedonic models. Repeat-sales price changes are unaffected by this problem as plot size hardly ever changes between transactions. Our method for estimating a proxy for plot size allows us to check specifically how big a problem missing plot size is for our hedonic results by computing K based on hedonic models that include and exclude our plot size proxy and then compare the results. We conjecture that missing plot size causes the hedonic model to underestimate the impact of improvements in energy efficiency on property prices.

The third bias is attenuation bias. If energy efficiency scores are measured with noise, attenuation bias will cause the hedonic model to underestimate the impact of improvements in energy efficiency on property prices Aydin et al. (2020).

The final bias is selection. This applies only to the repeat-sales price changes. More specifically, the improved energy efficiency properties are not randomly selected. In particular, properties with higher capitalisation rates of energy efficiency improvements are more likely to be selected into the repeat-sales with energy improvements sample. Selection bias again will tend to cause the price impact of energy improvements to be overestimated.

Our results for the K index are shown in Table 11. First, the measured price increases from the hedonic model are quite similar to those from the repeat-sales sample, differing on average by about 10%. This indicates that the hedonic regression model is generating plausible results. We find that K is less than 1 for all three property types. This is not surprising, given the various biases discussed above. Both the repeat-sales biases are upward, while two of the three hedonic biases are downward.

Including the plot size proxy in the hedonic model for detached houses slightly increases the K coefficient. This result aligns with our prior expectation as it somewhat reduces the downward bias in the hedonic predictions arising from the omitted plot size variable.

7.2 Using subsets of properties with more precisely measured quality to check for omitted-quality bias in the broader sample results

In the EPCs, we have more detailed information on the quality of the building structure of properties. More specifically, the EPCs provide information on the following: walls, roof, floor, windows, main heating, main heating controls, secondary heating, hot water and lighting. We repeat the calculation above for a different variant of the K coefficient. This variant, denoted by K^R, is defined below in (13). We now estimate the hedonic model on a restricted sample of properties – here, we focus on properties with double glazing. Let $\hat{p}^R_{t,n}(E_{t,n})$ denote the predicted price of property n in period t obtained from estimating the hedonic model on the restricted sample of properties with double glazing. Conversely, $\hat{p}_{t,n}(E_{t,n})$ denotes the predicted price of property n in period t obtained from the hedonic model estimated over the unrestricted sample. We now construct analogous versions of our K coefficients, except now the comparison is between restricted and unrestricted hedonic results rather than between hedonic and repeat-sales results.

$$K^R = \prod_{n=1}^{N_R} \left(\frac{\hat{p}^R_{t+k,n}(E_{t+k,n})}{\hat{p}^R_{t,n}(E_{t,n})} \bigg/ \frac{\hat{p}_{t+k,n}(E_{t+k,n})}{\hat{p}_{t,n}(E_{t,n})} \right)^{1/N_R}, \tag{13}$$

where $n = 1, \ldots, N_R$ denotes the sample of properties with double glazing.

By construction, the unobserved quality differences in the restricted sample with double glazing should be smaller than in the full sample. Hence the hedonic price predictions for the restricted sample obtained from the hedonic model estimated on the restricted sample should be less affected by omitted quality bias than those obtained from the hedonic model estimated on the full sample. Assuming that the omitted upward quality bias in predicted prices from energy improvements is smaller in the restricted hedonic model (as it should be), then it follows that K_R should be less than 1. Inspection of the K^R results in Table 11 reveals that this is indeed the case, although the effect is small.

As well as confirming our hypotheses regarding the biases in the hedonic results, Table 11 also suggests that these biases are quite small. Corrections that reduce the omitted variables bias have only very small effects on the predicted hedonic prices.

Table 11: Comparing Hedonic and Repeat-Sales Price Changes

Property type	K	Sample (K)	K^R	Sample (K^R)
Flats	0.925	1018	0.991	515306
Semi-Terraces	0.895	6246	0.994	2670895
Detached Houses (with plot proxy)	0.889	1252	0.996	751018

Note: The formulas for the K and K^R coefficients are provided in (12) and (13). “Sample” denotes the number of properties over which each coefficient is calculated. The samples cover transactions over the period 2008-2022.

8 Conclusion

Matching properties across EPCs and the HM Land Registry provides a rich, detailed micro-level dataset. We apply hedonic methods to this dataset to estimate the impact of the energy efficiency improvements recommended in EPCs on property prices. By comparing the price increases of these energy efficiency improvements with the associated costs provided in the EPCs, we were able to measure the Capitalisation Rate of Energy Efficiency Improvements (CREEIs).

Plot size is one key variable missing in both the EPC and HM Land Registry datasets. This characteristic is essential in a hedonic model for detached houses. We illustrate that its omission can cause the estimated coefficient on energy efficiency to be negative, particularly in areas with the highest land prices. Therefore, a proxy measure of plot size is needed to obtain meaningful results for detached houses, especially in London. We constructed this proxy plot size variable based on each property's distances to its ten nearest neighbours. A second issue with the data was that the costs of energy efficiency improvements in the EPCs had not been updated since 2011. We uprated the stated costs using an ONS price index for housing repairs and maintenance to reflect the price increases over the study period.

We find that, for almost all ITL-1 regions and property types over the period 2014-2022, the CREEI is less than 100, meaning that the cost of energy efficiency improvements are not fully capitalised in property prices. More specifically, for flats, the average CREEIs ranged between 72% (East) and 103% (East Midlands). For semis/terraces, they ranged between 47% (North East) and 65% (North West). For detached houses, they ranged between 24% (London) and 88% (North East). Hence the capitalisation rates are noticeably lower for semis/terraces and detached houses.

The CREEIs for detached houses in London are especially low. We hypothesise that this is due to the crucial role of location and plot size in London in determining property prices. Given the high land prices in London, these considerations tend to dominate other quality characteristics, such as energy efficiency. As shown in Table 3, the London region has by far the highest percentage of detached houses with energy efficiency level below C.

As a robustness check, we also computed the price impact of energy efficiency improvements from a sample of repeat-sale properties, where the energy efficiency was increased between sales. The repeat-sales results are similar to those obtained using our hedonic approach.

Housing is a major source of carbon emissions. Our main finding that energy efficiency improvements are not fully capitalised in property prices implies that owner incentives may need to be boosted if the UK is to achieve its Net Zero goal for carbon emissions by 2050. The "Help to Heat" scheme recently announced by the UK government (Department for Energy Security and Net Zero, 2023) may be useful in this regard, although it is important that such schemes are targeted to the regions and property types where they are most needed.

Finally, it is worth noting that the low energy efficiency ratings of detached houses in London raises the possibility of an awkward conflict between the government's twin agendas of reaching Net Zero by 2050 and Levelling Up. Levelling Up calls for subsidies to the poorest regions of the UK, while the Net Zero goal suggests providing subsidies to some owners of detached houses in London.

References

Asproudis, E., Gedikli, C., Talavera, O., and Yilmaz, O. (2023). Returns to solar panels in the housing market: A meta learner approach. Discussion Papers 23-01, Department of Economics, University of Birmingham.

Aydin, E., Brounen, D., and Kok, N. (2020). The capitalization of energy efficiency: Evidence from the housing market. Journal of Urban Economics, 117:103243.

Bostic, R. W., Longhofer, S. D., and Redfearn, C. L. (2007). Land leverage: Decomposing home price dynamics. Real Estate Economics, 35(2):183–208.

Brounen, D. and Kok, N. (2011). On the economics of energy labels in the housing market. Journal of Environmental Economics and Management, 62(2):166–179.

Bruegge, C., Carrión-Flores, C., and Pope, J. C. (2016). Does the housing market value energy efficient homes? Evidence from the energy star program. Regional Science and Urban Economics, 57:63–76.

Cerin, P., Hassel, L. G., and Semenova, N. (2014). Energy performance and housing prices: Does higher dwelling energy performance contribute to price premiums? Sustainable Development, 22(6):404–419.

Cook, R. D. (1979). Influential observations in linear regression. Journal of the American Statistical Association, 74:169–174.

de Ayala, A., Galarraga, I., and Spadaro, J. V. (2016). The price of energy efficiency in the Spanish housing market. Energy Policy, 94:16–23.

Deng, Y., Li, Z., and Quigley, J. M. (2012). Economic returns to energy-efficient investments in the housing market: Evidence from Singapore. Regional Science and Urban Economics, 42(3):506–515.

Department for Business, Energy and Industrial Strategy (2022). Mission zero: Independent review of net zero. Final Report, 26 September, 2022. https://www.gov.uk/government/publications/review-of-net-zero.

Department for Energy Security and Net Zero (2023). Find energy grants for your home. 23 March, 2023. https://www.gov.uk/government/collections/find-energy-grants-for-you-home-help-to-heat.

Department for Levelling Up, Housing and Communities (2022). Energy performance of buildings data: England and Wales. 11 November, 2022. https://epc.opendatacommunities.org/docs/guidance.

Department for Levelling Up, Housing and Communities (2023). Energy performance of buildings data England and Wales - guidance. January 20, 2023, https://epc.opendatacommunities.org/docs/guidance.

Diewert, W. E. (2009). Durables and owner-occupied housing in a consumer price index. Price Index Concepts and Measurement, edited by W. E. Diewert, J. S. Greenlees and C. R. Hulten, pages 445–500. University of Chicago Press.

Dinan, T. M. and Miranowski, J. A. (1989). Estimating the implicit price of energy efficiency improvements in the residential housing market: A hedonic approach. Journal of Urban Economics, 25(1):52–67.

Eichholtz, P., Kok, N., and Quigley, J. M. (2010). Doing well by doing good? Green office buildings. American Economic Review, 100(5):2492–2509.

Eurostat, Organisation, I. L., Fund, I. M., Organisation for Economic Co-operation and Development, for Europe, U. N. E. C., and the World Bank (2013). Handbook on residential property prices indexes (RPPIs). European Union, Luxembourg, Chapter 5 on Hedonic Regression Methods by J. de Haan and W. E. Diewert.

Evangelista, R., Ramalho, E. A., and Andrade e Silva, J. (2020). On the use of hedonic regression models to measure the effect of energy efficiency on residential property transaction prices: Evidence for Portugal and selected data issues. Energy Economics, 86(C):104699.

Evangelista, R., Ramalho, E. A., and Andrade e Silva, J. (2022). How heterogeneous is the impact of energy efficiency on dwelling prices? Evidence from the application of the unconditional quantile hedonic model to the Portuguese residential market. Energy Economics, 109:105955.

Fuerst, F. and McAllister, P. (2011). Green noise or green value? Measuring the price effects of environmental certification in commercial buildings. Real Estate Economics, 39(1):45–69.

Fuerst, F., McAllister, P., Nanda, A., and Wyatt, P. (2015). Does energy efficiency matter to home-buyers? An investigation of EPC ratings and transaction prices in England. Energy Economics, 48(C):145–156.

Fuerst, F., McAllister, P., Nanda, A., and Wyatt, P. (2016). Energy performance ratings and house prices in Wales: an empirical study. Energy Policy, 92:22–33.

Fuerst, F. and Shimizu, C. (2016). Green luxury goods? The economics of eco-labels in the Japanese housing market. Journal of the Japanese and International Economies, 39:108–122.

Fuerst, F. and Warren-Myers, G. (2018). Does voluntary disclosure create a green lemon problem? Energy-efficiency ratings and house prices. Energy Economics, 74:1–12.

Geske, J. (2022). The value of energy efficiency in residential buildings – a matter of heterogeneity?! Energy Economics, 113:106173.

Hall, R. E. and Jorgenson, D. W. (1967). Tax policy and investment behavior. American Economic Review, 57(3):391–414.

Hill, R. and Melser, D. (2008). Hedonic imputation and the price index problem: an application to housing. Economic Inquiry, 46(4):593–609.

Hill, R. J. and Syed, I. A. (2016). Hedonic price-rent ratios, user cost, and departures from equilibrium in the housing market. Regional Science and Urban Economics, 56:60–72.

HM Land Registry (2023). Price-paid-data. 10 February, 2023. https://www.gov.uk/government/statistical-data-sets/price-paid-data-downloads.

Hyland, M., Lyons, R., and Lyons, S. (2013). The value of domestic building energy efficiency – Evidence from Ireland. Energy Economics, 40:943–952.

Hyslop, R. and Imbens, G. W. (2001). Bias From Classical and Other Forms of Measurement Error. Journal of Business and Economic Statistics, 19(4):475–481.

Johnson, R. C. and Kaserman, D. L. (1983). Housing market capitalization of energy saving durable good investments. Economic Inquiry, 21:374–386.

Kahn, M. E. and Kok, N. (2014). The capitalization of green labels in the California housing market. Regional Science and Urban Economics, 47:25–34.

Kholodilin, K. A., Michelsen, C., and Mense, A. (2017). The market value of energy efficiency in buildings and the mode of tenure. Urban Studies, 54(14):3218–3238.

Kok, N., McGraw, M., and Quigley, J. M. (2011). Housing market capitalization of energy saving durable good investments. American Economic Review, 101(2):77–82.

Lancaster, K. J. (1966). A new approach to consumer theory. Journal of Political Economy, 74:132–157.

Laquatra, J. (1986). Housing market capitalization of thermal energy. Energy Economics, 8(3):134–138.

Marmolejo-Duarte, C. and Chen, A. (2022). The effect of energy performance ratings over residential prices or how an insufficient control of architectural-quality may render spurious conclusions. Cities, 126:103674.

Ministry of Housing, Communities and Local Government (2019). National statistics, english indices of deprivation 2019. 10 February, 2023. https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019.

Office for National Statistics (2021). Energy efficiency of housing in England and Wales. 10 November, 2021. https://www.ons.gov.uk/peoplepopulationandcommunity/housing/articles/energyefficiencyofhousinginenglandandwales/2021.

Office for National Statistics (2023). About the area classifications: Information and frequently asked questions about the 2011 area classifications. 10 February, 2023. https://www.ons.gov.uk/methodology/geography/geographicalproducts/areaclassifications/2011areaclassifications/abouttheareaclassifications.

Poterba, J. M. (1984). Tax subsidies to owner-occupied housing: an asset-market approach. Quarterly Journal of Economics, 99(4):729–752.

Pride, D., Little, J., and Mueller-Stoffels, M. (2018). The value of residential energy efficiency in interior Alaska: A hedonic pricing analysis. Energy Policy, 123:450–460.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. Journal of Political Economy, 82(1):34–55.

Rousseeuw, P. J. and Leroy, A. M. (1987). Robust regression and outlier detection. Wiley, New York.

Silver, M. (2018). How to measure hedonic property price indexes better. EURONA – Eurostat Review on National Accounts and Macroeconomic Indicators, 2018/1:article 2.

Silver, M. (2022). Econometric Issues in Hedonic Property Price Indices: Some Practical Help. Journal of Official Statistics, 38:153–186.

S&P Dow Jones Indices (2022). S&P Corelogic Case-Shiller Home Price Indices Methodology. S&P Dow Jones Indices: Index Methodology, 10 April, 2022. https://www.spglobal.com/spdji/en/documents/methodologies/methodology-sp-corelogic-cs-home-price-indices.pdf.

Taruttis, L. and Weber, C. (2022). Estimating the impact of energy efficiency on housing prices in Germany: Does regional disparity matter? Energy Economics, 105:105750.

UK Government (2023a). Buying or selling your home. 10 February, 2023. https://www.gov.uk/buy-sell-your-home/energy-performance-certificates.

UK Government (2023b). Guidance on price-paid-data. 10 February, 2023. https://www.gov.uk/guidance/about-the-price-paid-data#data-excluded-from-price-paid-data.

UK Government (2023c). Property boundaries. 10 February, 2023. https://www.gov.uk/your-property-boundaries.

Wahlström, M. H. (2016). Doing good but not that well? a dilemma for energy conserving homeowners. Energy Economics, 60:197–205.

Walls, M., T., Gerarden, K., Palmer, and Bak, X. F. (2017). Is energy efficiency capitalized into home prices? Evidence from three U.S. cities. Journal of Environmental Economics and Management, 82:104–124.

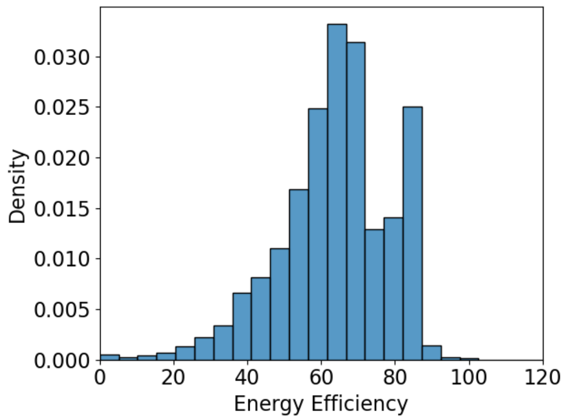
Welsh Government (2019). Welsh index of multiple deprivation (full index update with ranks): 2019. 10 February, 2023. https://www.gov.wales/welsh-index-multiple-deprivation-full-index-update-ranks-2019.

Zhang, L., Liu, H., and Wu, J. (2017). The price premium for green-labelled housing: Evidence from China. Urban Studies, 54(15):3524–3541.

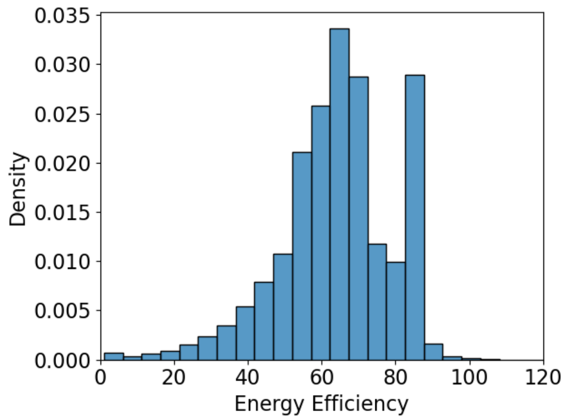
Appendix A: Distribution of Energy Efficiency

Figure A1: Energy Efficiency: merged transaction data (left), merged EPC data (right)

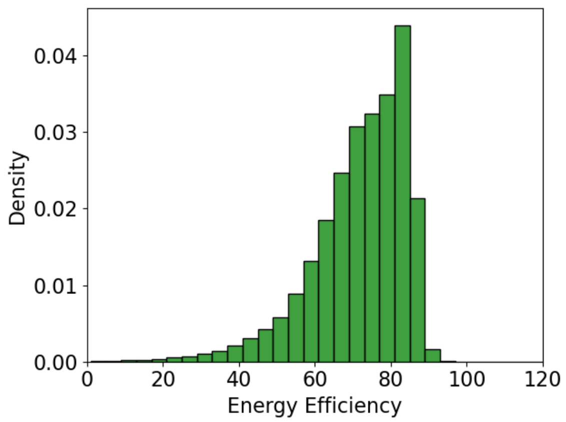
(a) Energy Efficiency for Detached Houses: 1 510 885 observations



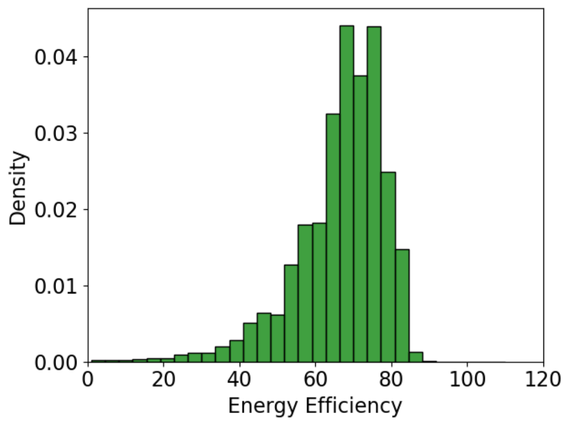
(b) Energy Efficiency for Detached Houses: 1 335 157 observations



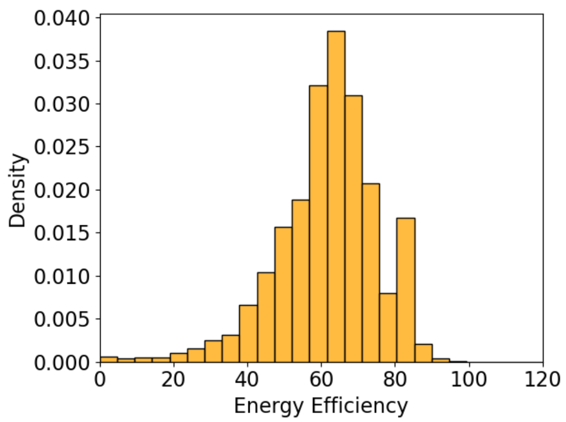
(c) Energy Efficiency for Flats: 730 579 observations



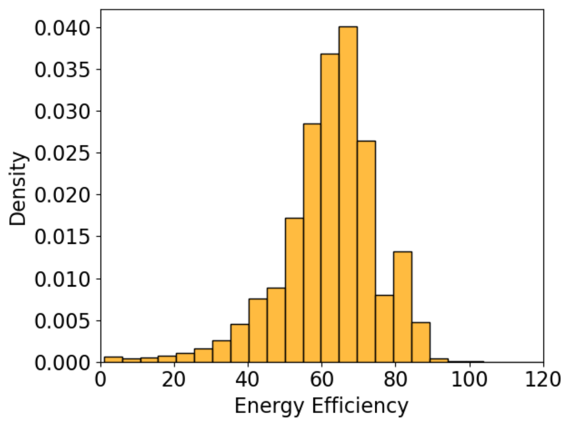
(d) Energy Efficiency for Flats: 559 200 observations



(e) Energy Efficiency for Semis/Terraces: 3 397 605 observations



(f) Energy Efficiency for Semis/Terraces: 3 200 744 observations



Note: The left panel shows the distribution of current energy efficiency for the data used to construct the hedonic regression models, whereas the right panel shows the distribution of the current energy efficiency for the data used for the evaluation (all EPCs).

Appendix B: Results without Uprating Recommendation Costs

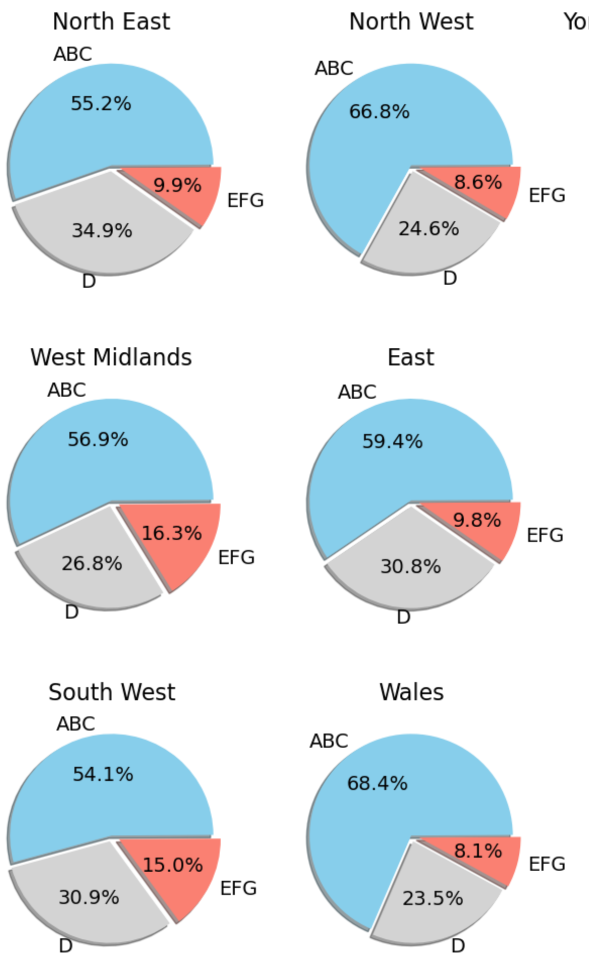
Table B1: Capitalisation Rate of Energy Improvements (CREEI) - with Costs Not Uprated

	North East	North West	Yorkshire and the Humber	East Midlands	West Midlands	East	London	South East	South West	Wales
Flats	$\boxed{92.0}$	101.1	84.6	108.5	84.2	76.1	95.6	77.4	79.7	87.7
Semi Terrace Properties	49.6	68.6	66.8	65.3	63.6	60.1	68.2	63.8	54.0	66.7
Detached Houses	74.8	40.5	41.1	53.2	38.0	45.2	-20.7	9.3	33.4	43.6
Detached Houses with plot size	93.0	68.5	63.2	68.3	64.9	67.5	25.3	59.6	57.8	58.1

Note: The formula for computing the CREEI is provided in (11). The CREEI for flats located in the North East is on average $\boxed{92.0\%}$. This implies that when all recommendations in the EPC are implemented, the price of a flat in the North East rises on average at a rate of 92 Pounds for every 100 Pounds of expenditure on energy efficiency improvements.

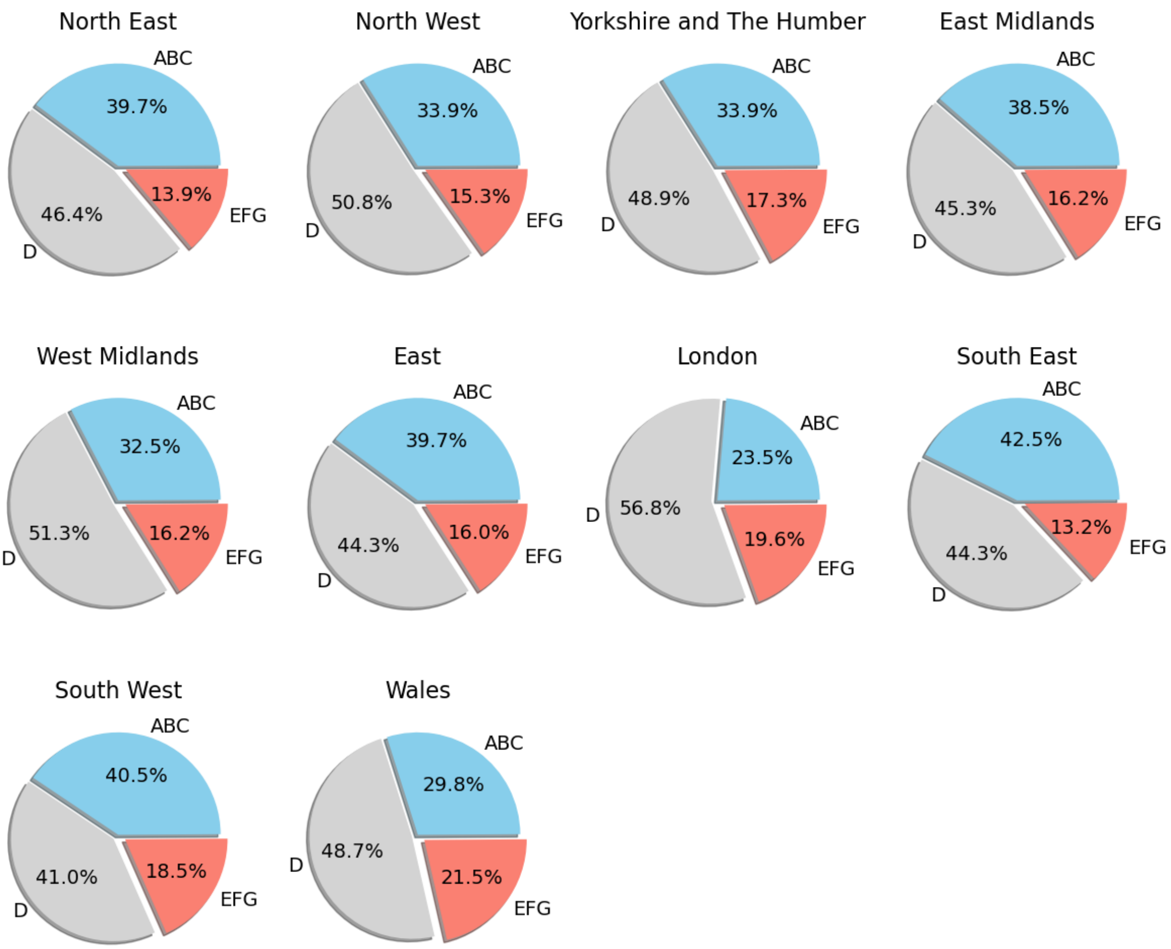
Appendix C: Regional Differences in Energy Efficiency Class Across England and Wales

Figure C1: Energy Efficiency Class per Region, Flats (2019)



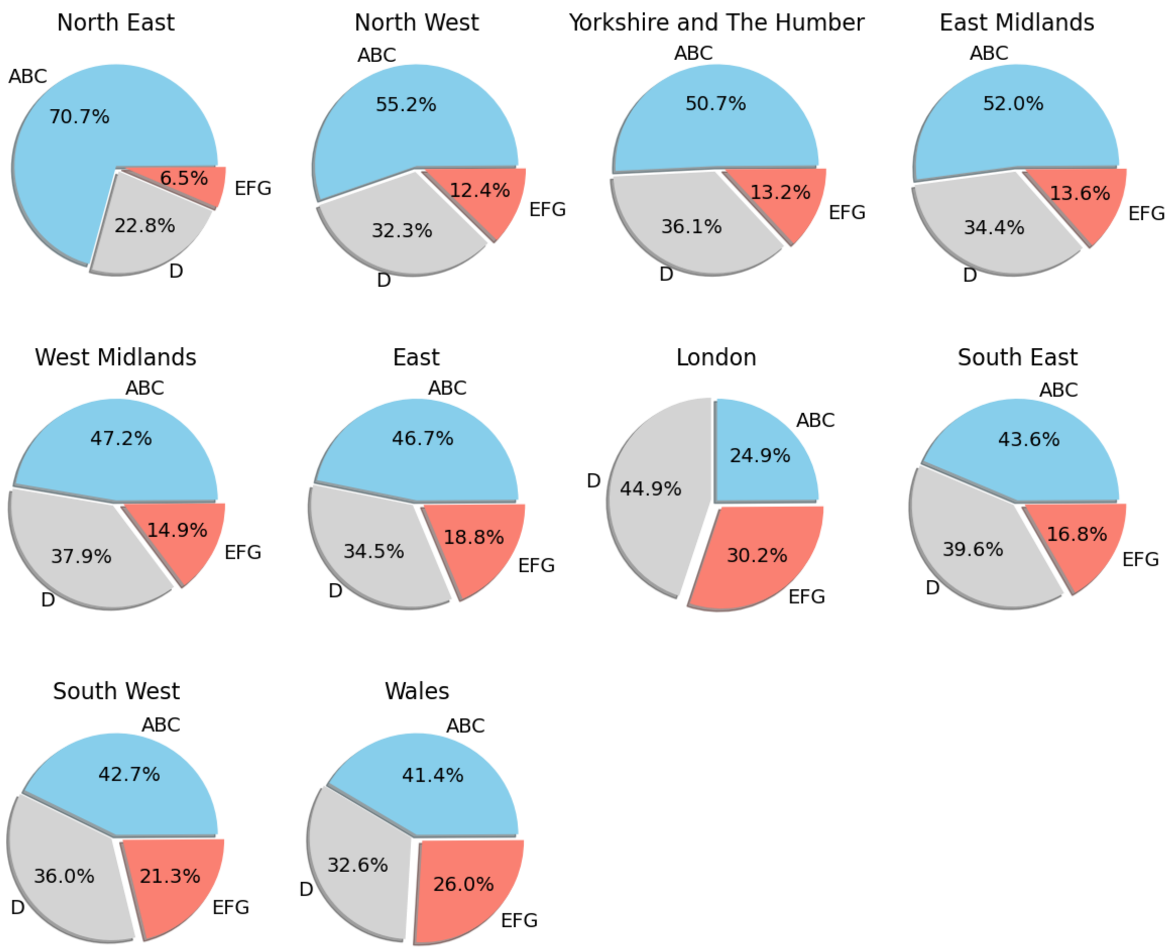
Note: Figure C1 corresponds to all transactions for flats in 2019 that could be matched with a EPC not older than 2 years.

Figure C2: Energy Efficiency Class per Region, Semi-Terrace (2019)



Note: Figure C2 corresponds to all transactions for semi-terraced houses in 2019 that could be matched with a EPC not older than 2 years.

Figure C3: Energy Efficiency Class per Region, Detached Houses (2019)



Note: Figure C3 corresponds to all transactions for detached houses in 2019 that could be matched with a EPC not older than 2 years.

Appendix D: Results without cleaning

Table D1 show the results from Table 10, but with the regression model estimated without any cleaning process.

Table D1: Flats: Capitalisation Rate of Energy Improvements (CREEI) without Data Cleaning

	North East	North West	Yorkshire and the Humber	East Midlands	West Midlands	East	London	South East	South West	Wales
Flats	$\boxed{82.1}$	78.9	87.5	118.7	96.0	95.0	77.3	86.7	83.6	87.7

Note: The formula for computing the CREEI is provided in (11). The CREEI for the North East is on average $\boxed{82.1\%}$. This implies that when all recommendations in the EPC are implemented, the price of a flat in the North East rises on average at a rate of 82.1 Pounds for each 100 Pounds of expenditure on energy efficiency improvements.

Appendix E: Matching Rates

Table E1: Matching rates for properties in the PPD ad EPC datasets in 2019

	Matched	Unmatched	All
Number of observations	580 565	81 085	661 650
Share of total transactions	88 %	12 %	100 %
Share new properties	54%	46%	100 %
Share freehold	90 %	10%	100 %
Share ppd_cat (market transactions)	91 %	9%	100 %
Mean price	£348 138	£434 998	£357 099
Standard deviation price	£367 522	£518 444	£971 755

Note: The 12% unmatched properties in 2019 include commercial transactions that can not be linked with residential EPCs and transactions with address errors in either the PPD or EPC datasets. Properties designated as "new" account for about 2% of the transaction dataset.

Appendix F: Non-Linearity Check of Energy Efficiency Coefficient

Table F1: Regression results for interaction term of energy efficiency score and rating class

Energy rating	Flats	Semi-Terraced	Detached House
A	0.0017***	0.0029***	0.0018***
B	0.0023***	0.0031***	0.0019***
C	0.0022***	0.0029***	0.0015***
D	0.0023***	0.0031***	0.0017***
E	0.0021***	0.0033***	0.0020***
F	0.0022***	0.0031***	0.0021***
G	0.0011*	0.0015***	0.000019

Note: Table F1 shows the estimated result for the energy efficiency coefficient, when using a separate coefficient for each energy rating class. Asterisks representing significance. * p-value <= 0.05; ** p-value <= 0.01, *** p-value <= 0.001.

Appendix G: Graphs of bootstrapped standard errors on the average CREEIs

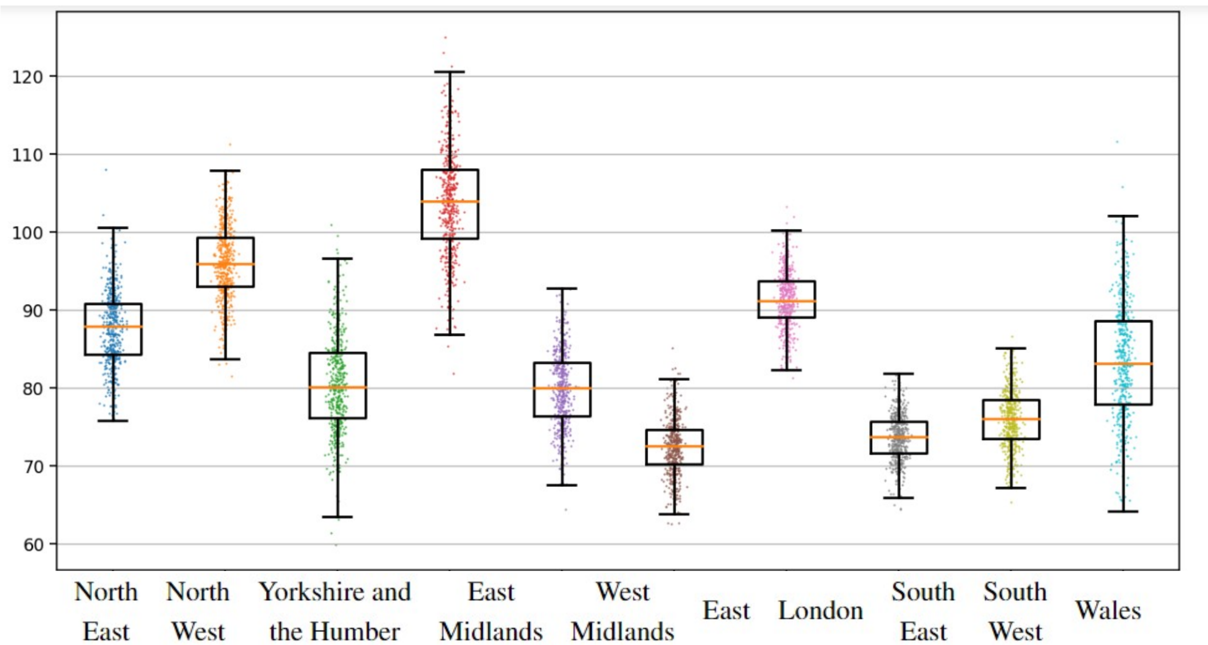


Figure G1: Bootstrap Results of Average CREEI for flats using 500 rounds

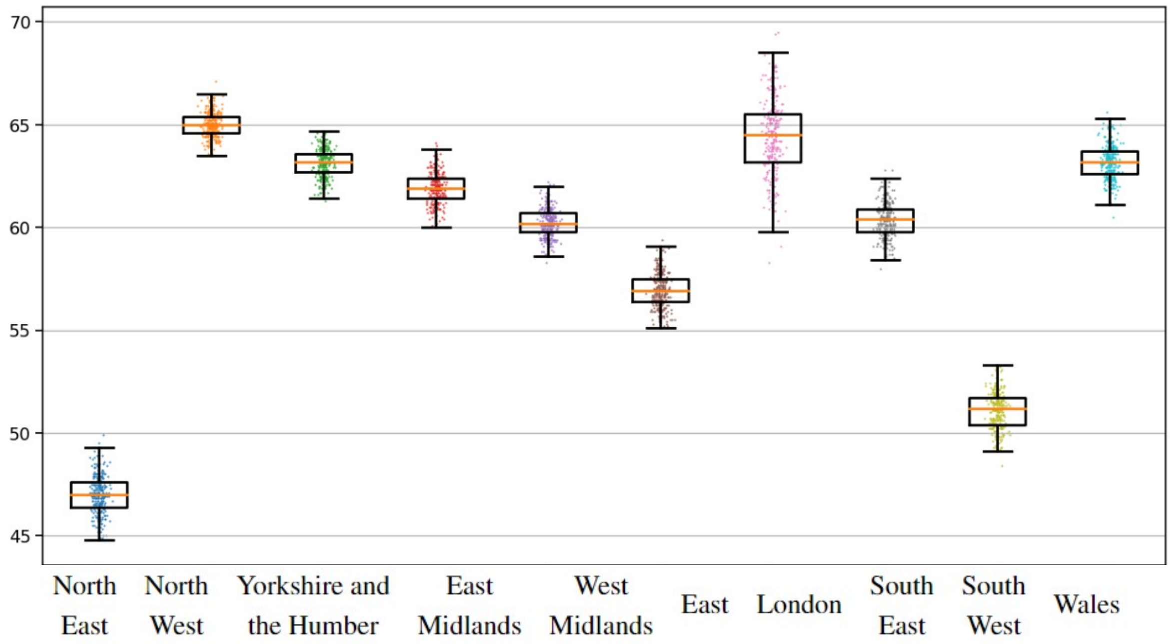


Figure G2: Bootstrap Results of Average CREEI for semi-terraced properties using 300 rounds

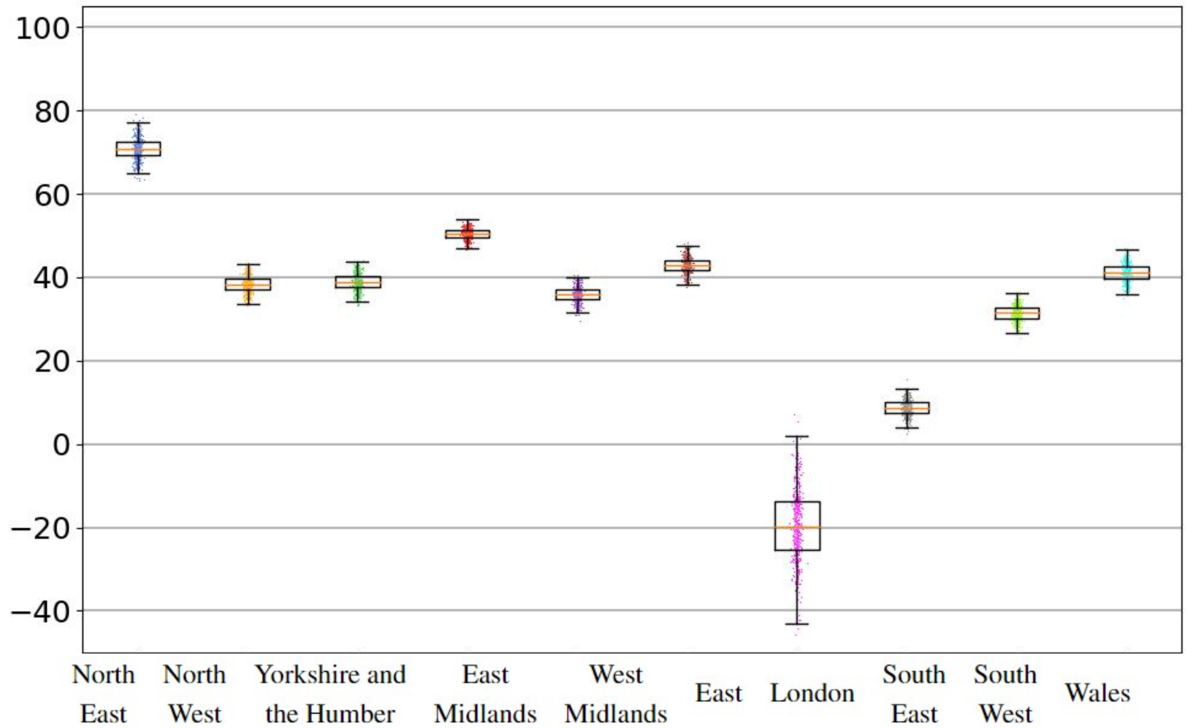


Figure G3: Bootstrap Results of Average CREEI for detached houses excluding plot-size using 500 rounds

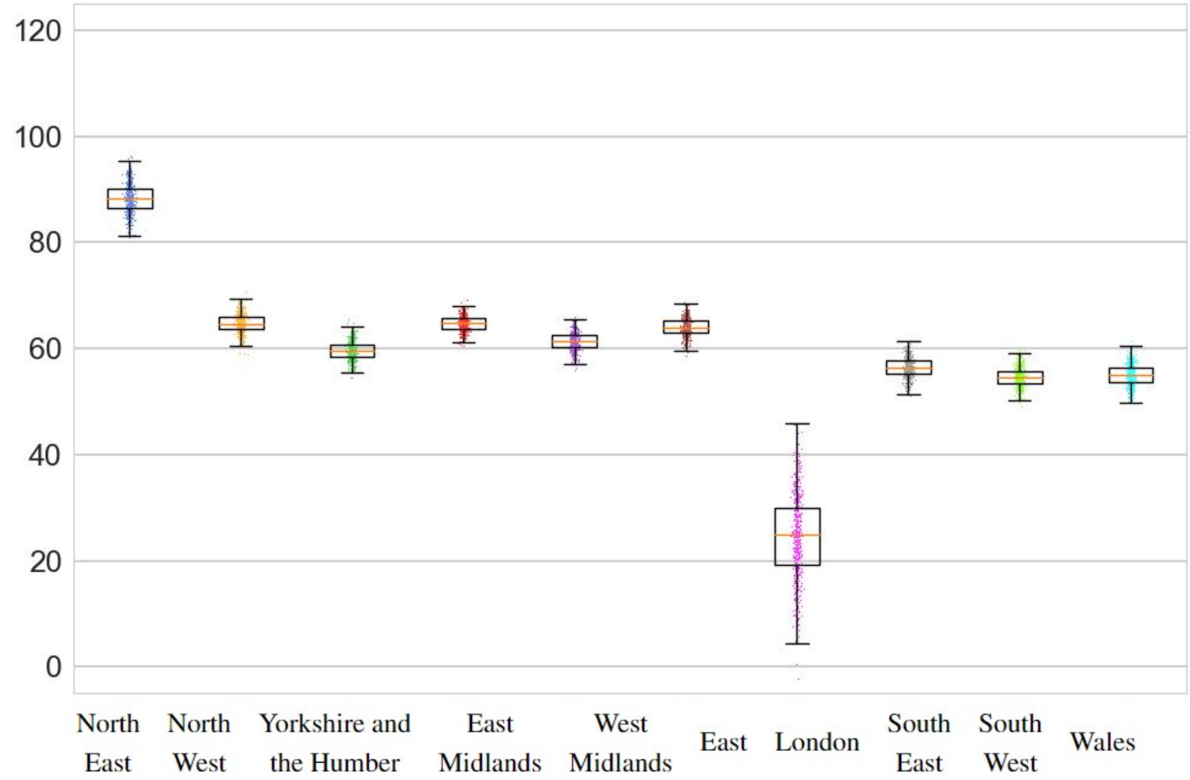


Figure G4: Bootstrap Results of Average CREEI for detached houses including plot-size using 500 rounds

Appendix H: Measuring market incentives for below-level-C properties

We extend the analysis here to focus specifically on properties that are initially below level C. The reason is that these properties have the greatest potential for increasing energy efficiency. The distribution of properties below level C is shown in Table 3. What is particularly striking in this Table is the high percentage of detached houses in London that are below level C.

The shares of potential energy classes of properties that start below level C if they implemented all the recommendations in the EPCs are shown in Table H1. As can be seen, over 83% of flats and 90% of semis/terraces and detached houses would reach level C if all recommended energy efficiency improvements were implemented.

Table H1: Distribution of potential energy efficiency for properties that have a current energy efficiency below level C, all EPCs

Energy Efficiency Rating		Flats	Semi-Terrace	Detached
Level C or above	A	0.00%	1.72%	2.81%
	B	7.06%	63.67%	46.86%
	C	76.82%	29.47%	45.06%
Below Level C	D	14.67%	4.3%	4.62%
	E	1.25%	0.66%	0.53%
	F	0.17%	0.15%	0.10%
	G	0.03%	0.03%	0.02%
	Total	100%	100%	100%

Table H1 shows what the energy efficiency rating label would be for properties starting below level C if they implemented all the recommendations in the EPC.

The CREEIs for properties starting below level C are shown Table H2.[23] The results for flats in Table H2 are similar to those in Table 10. However, for below level C semis/terraces, the average CREEIs are higher for 8 out of 10 regions, while for all 10 regions the CREEIS are higher for below level C detached houses. The implication is that for semi-terraces and houses, the market incentives to increase energy efficiency are a bit higher for the least energy-efficient properties (although the CREEIs are generally still well below 100%).

Table H2: Capitalisation Rate of Energy Efficiency Improvements (CREEI) for Properties Initially Below Level C

	North East	North West	Yorkshire and the Humber	East Midlands	West Midlands	East	London	South East	South West	Wales
Flats	$\boxed{82.8}$	103.4	78.6	92.5	84.3	70.1	89.3	75.8	76.2	73.0
Semi-Terraced Properties	48.1	67.1	64.4	61.5	61.4	60.4	67.3	64.6	54.6	62.9
Detached Houses with plot size	104.2	75.6	68.6	69.5	67.0	70.7	25.5	63.4	62.0	59.9

Note: The formula for computing the CREEI is provided in (11). The average CREEI for Flats locateed in the North East is $\boxed{82.8\%}$. This implies that when all recommendations in the EPC are implemented, the price of a flat in the North East rises on average at a rate of 82.8 Pounds for every 100 Pounds of expenditure on energy efficiency improvements.

[23] Results for detached houses without plot size are not shown.

Appendix I: Regression Summary of Flats and Semi-Terrace Properties

Table I1: Regression Output for Hedonic Models for Flats

	North East	North West	Yorkshire and the Humber	East Midlands	West Midlands	East	London	South East	South West	Wales
Constant	9.8597***	10.0561***	10.2467***	10.3330***	10.5380***	10.4678***	12.0974***	10.7745***	10.8257***	10.9951***
	(0.052)	(0.035)	(0.032)	(0.034)	(0.027)	(0.018)	(0.069)	(0.014)	(0.020)	(0.040)
Energy Efficiency	0.0044***	0.0027***	0.0021***	0.0031***	0.0020***	0.0016***	0.0014***	0.0015***	0.0020***	0.0025***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Floor Area	0.0127***	0.0132***	0.0124***	0.0119***	0.0122***	0.0114***	0.0116***	0.0117***	0.0130***	0.0118***
	(0.009)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Deprivation	1.435e-05***	9.689e-06***	1.225e-05***	8.098e-06***	1.122e-05***	6.886e-06***	9.687e-06***	6.154e-06***	8.77e-06***	0.0001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Flat	Base	Base	Base	Base	Base	Base	Base	Base	Base	Base
Maisonette	-0.0297***	-0.0671***	-0.0709***	-0.0163***	-0.0168***	0.0000	-0.0124***	0.0025	-0.0450***	-0.0052
	(0.009)	(0.007)	(0.007)	(0.006)	(0.004)	(0.003)	(0.002)	(0.002)	(0.004)	(0.009)
Freehold	Base	Base	Base	Base	Base	Base	Base	Base	Base	Base
Leasehold	-0.0312	0.0430	0.0602***	-0.0708***	-0.0460***	-0.0082	-0.0471	0.0048	0.0500***	-0.0059
	(0.017)	(0.023)	(0.010)	(0.014)	(0.014)	(0.008)	(0.027)	(0.006)	(0.007)	(0.019)
Supergroup	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Construction Age	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tenure	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Built Form	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Local Authority	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16,472	50,565	29,703	19,206	41,410	80,313	259,303	147,163	72,220	14,224
Adjusted R^2	0.68	0.71	0.70	0.66	0.76	0.82	0.84	0.78	0.72	0.70

The table presents the coefficients of the energy efficiency parameter for flats. The dependent variable is the logarithm of the transaction price. The model is run separately for each ITL-1 region with data from 2008-2022. The characteristics include: Floor Area: in square meters; Deprivation: a rank index on relative deprivation in small areas in England; Supergroup levels: “Rural”, “Cosmopolitans”, “Ethnicity”, “Multicultural”, “Urbanites”, “Suburbanites”, “Constrained”, “Hard-pressed”; Property Type levels: Flat, Maisonette; Construction Age levels: before 1900, 1900-1929, 1930-1949, 1950-1966, 1967-1975, 1976-1982, 1983-1990, 1991-1995, 1996-2002, 2003-2006, 2007 onward; Property Type levels: Flat, Maisonette; Tenure levels: Owner-Occupied, Rental Private, Rental Social; Duration levels: Leasehold, Freehold; Built Form levels: Detached Semi-Detached, End-Terrace, Mid-Terrace; and Year levels: 2008 to 2022. It should be noted that freehold flats are considered more risky than leasehold, due to only 0.9% of flats in our sample are freehold. Also, the quality of freehold flats varies by region.
Note: Asterisks representing significance. * p-value <= 0.05; ** p-value <= 0.01, *** p-value <= 0.001.

Table I2: Regression Output for Semi/Terrace Properties

	North East	North West	Yorkshire and the Humber	East Midlands	West Midlands	East	London	South East	South West	Wales
Constant	10.6053***	10.6746***	10.6088***	10.8827***	11.0404***	11.0999***	12.3731***	11.4250***	11.4866***	10.6185***
	(0.013)	(0.008)	(0.009)	(0.008)	(0.012)	(0.007)	(0.053)	(0.006)	(0.008)	(0.014)
Energy Efficiency	0.0037***	0.0040***	0.0039***	0.0036***	0.0033***	0.0020***	0.0012***	0.0018***	0.0020***	0.0043***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Floor Area	0.0084***	0.0075***	0.0075***	0.0068***	0.0071***	0.0062***	0.0067***	0.0064***	0.0066***	0.0070***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Deprivation	1.635e-05***	1.541e-05***	1.572e-05***	1.011e-05***	1.202e-05***	7.56e-06***	1.267e-05***	7.779e-06***	9.573e-06***	0.0002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
House	Base	Base	Base	Base	Base	Base	Base	Base	Base	Base
Bungalow	0.1341***	0.0737***	0.0676***	0.0775***	0.0922***	0.0566***	0.0277***	0.0512***	0.0837***	0.0938***
	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.003)
Flat	-0.0992***	-0.0169**	-0.0281***	0.0057	-0.0199***	-0.0256***	-0.0194***	0.0058	-0.0334***	-0.0768***
	(0.008)	(0.006)	(0.006)	(0.001)	(0.006)	(0.004)	(0.004)	(0.003)	(0.004)	(0.008)
Maisonette	-0.0319	-0.0781***	-0.0818***	-0.0762***	-0.1685***	-0.0636***	-0.0797***	-0.0508***	-0.0536***	-0.0557**
	(0.020)	(0.014)	(0.012)	(0.011)	(0.010)	(0.005)	(0.003)	(0.005)	(0.008)	(0.003)
Freehold	Base	Base	Base	Base	Base	Base	Base	Base	Base	Base
Leasehold	0.0033	-0.0174**	0.0361***	-0.1047***	-0.0367***	-0.2090***	-0.1451***	-0.2151***	-0.0589***	-0.0168***
	(0.002)	(0.001)	(0.002)	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Supergroup	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Construction Age	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tenure	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Built Form	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Local Authority	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	161,370	457,487	354,114	286,954	337,226	390,429	338,033	550,709	354,131	167,152
Adjusted R^2	0.75	0.76	0.76	0.75	0.77	0.87	0.88	0.84	0.73	0.74

The table presents the coefficients of the energy efficiency parameter for semi-detached and terraced houses. The dependent variable is the logarithm of the transaction price. The model is run separately for each ITL-1 region with data from 2008-2022. The characteristics are as in Table I1 above. All included characteristics have the expected sign and are highly significant. Note: Asterisks representing significance. * p-value <= 0.05; ** p-value <= 0.01, *** p-value <= 0.001.