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

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

February 27, 2025

Abstract

Increasing the energy efficiency of housing is a key part of strategies to reduce carbon emissions. In this paper we estimate a hedonic model to measure the market incentives of owners to improve the energy efficiency of residential properties in England and how these incentives differ by location and property type. To do this, we merge sales records with Energy Performance Certificates (EPCs) to create a micro-level dataset consisting of 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. Our results reveal significant differences in market incentives across regions and property types. On average, we find that 52% of the costs of EPC-recommended energy efficiency improvements are capitalized in property prices for flats, as compared with 60% for semis/terraces and 23% for detached houses. For London only 2.5% of energy efficiency improvements for detached houses are capitalized. Checks based on regression discontinuity design, two-stage least squares and repeat sales confirm the robustness of our baseline results. (JEL. C43; Q52; Q54; R31)

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

Competing interests statement: None of the authors have any competing interests.

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), the Wegener Center seminar (June 15, 2023), University of New South Wales (December 13, 2023), and Cambridge University (October 23, 2024) 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 set a goal for the UK to reach net zero carbon emissions by 2050 (Department for Business, Energy and Industrial Strategy, 2022). An important part of this strategy is to find ways of increasing the energy efficiency of residential housing. In this paper we assess the existing 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 needed to design an effective Net-Zero strategy.

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, especially when mandatory, 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 (Myers et al., 2022). This, in turn, should incentivize 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 energy 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 and comparing them with the associated cost of these improvements. To this end, we match the records for individual residential property transactions in England from the HM Land Registry with EPCs over the period 2015-2023. We then use a hedonic approach to estimate the impact of 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) for the regions All England, London and the Rest of England. Estimates of the costs of the recommended energy efficiency improvements are provided directly in the EPCs.

We calculate the capitalization 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 capitalized 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 additional property characteristics and features.

Plot size, however, is a key variable that is missing from both datasets.[1] We construct a proxy measure of plot size using an innovative approach based on the distances between neighboring 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 and government properties. Our plot size indicator variable is highly significant when included in our hedonic models.

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 three checks for bias in our hedonic results. First, we undertake a regression discontinuity design (RDD) analysis of price jumps at label boundaries. Second, we estimate a two-stage least squares (2SLS) version of our hedonic model using the energy price as an instrument. Third, we compare our hedonic results with results obtained from an analysis of repeat sales. The RDD analysis does not generate significant price jumps at the label boundaries and hence is not useful in our context. The 2SLS analysis generates results that are very similar to those obtained from our baseline ordinary least squares (OLS) hedonic analysis. The repeat-sales analysis generates price increases from energy efficiency improvements that are between 10% and 16% larger (depending on the property type) than those obtained using

[1] The plot size is not usually recorded in English 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.

OLS or 2SLS hedonic approaches. It should be noted though that the repeat-sales results are based on a small sub-sample of the dataset used in the hedonic analysis. The checks provided by the 2SLS and repeat-sales analyses suggest that our results are reasonably robust.

Turning to the issue of market incentives over the period 2015-2023, we find that tbe CREEI is 59% for flats, 63% for semis/terraces and 21% for detached houses. A CREEI of 59% for flats, for example, implies that 59% of the costs of energy efficiency improvements are capitalized in higher flat prices. The corresponding CREEIs for London are higher for flats and semis/terraces but lower for detached houses, where the CREEI is only 2.5%.

We consider possible explanations for the very low observed CREEI for detached houses in London. To the extent this estimate is genuine, it creates a difficult policy dilemma. Subsidies and other support to reduce carbon emissions in the housing sector will be most effective if they are targeted to regions and property types where market incentives are weakest (i.e., detached houses in London). On the other hand, owners of detached houses in London tend to be much more wealthy than the average household in England. Providing these households with subsidies or other incentives may therefore increase inequality.

The remainder of this paper is structured as follows. Section 2 discusses some of the relevant existing literature. Section 3 outlines a hedonic model of property prices that includes energy efficiency. The dataset is described in section 4. The results and robustness checks are presented in section 5 and the conclusion is in section 6.

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).

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. (2024) 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. (2024) 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 than for detached and semis/terraces. 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 2008 (and have remained mandatory in England after Brexit).[2]

[2]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.

Some authors caution that the price premia 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. Sejas-Portillo et al. (2024) use a regression discontinuity design (RDD) applied to UK data to the label boundaries to identify the pure price effect of energy efficiency improvements while Aydin et al. (2020) use two stage least squares (2SLS) applied to Dutch data with the instrument being a dummy for properties built before or after the 1974 oil price shock. In our empirical analysis we assess the feasibility of RDD, 2SLS and repeat-sales as alternatives for measuring the price impact of energy efficiency improvements.

3 Theoretical Model

The theoretical model presented here treats housing H as a differentiated good:

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

where Q refers to observed housing characteristics, such as the property's floor area, age and locational aspects. Q^o refers to omitted characteristics that are not available in the dataset (such as plot size, the quality of the building materials, kitchen, and other fittings).

Suppose now that households have the utility function

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

where we treat the initial level of housing quality, $\bar{H}$, as given. The energy efficiency level is also given and denoted by $\bar{E}$. A household must decide how much to increase the energy efficiency of the property, ΔE, and how much to spend on a composite consumption good, x.

The cost of owner-occupying can be expressed as a user cost (Poterba, 1984). In equilibrium, individuals should be indifferent between owning and renting such that the yearly 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.[3]

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

$$uP(\bar{H}, \bar{E}) + x + C(\Delta E) = M + P(\bar{H}, \bar{E} + \Delta E) - P(\bar{H}, \bar{E}), \tag{5}$$

where price (P) is assumed to be a function of the observed and unobserved characteristics (H) and energy efficiency (E), while $C(\Delta E)$ is the cost of the energy efficiency improvement, ΔE. The price of the composite consumption good, x, is normalised to 1.

Households choose x, and ΔE to maximise utility subject to their budget constraint. Solving for x and ΔE from the first-order conditions, it follows that the equilibrium property price, P, is a function of the following variables and parameters:

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

where c represent the parameters from the cost function, C, 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 the price of a property associated with the change in its energy efficiency from $\bar{E}$ to $\bar{E} + \Delta E$ while holding all other variables fixed: $P(\bar{E} + \Delta E) - P(\bar{E})$. We then compare $P(\bar{E} + \Delta E) - P(\bar{E})$ with $C(\Delta E)$ at the level of individual properties.

The problem is that we cannot directly control for omitted variables, Q^o. The better the dataset is able to describe the property (i.e., the more price-relevant variables are in Q rather than Q^o), the smaller this omitted variable bias should get. One aspect of this paper is to transform plot size, an important price-determining characteristic variable, from a type Q^o into a type Q variable. Another aspect is to perform robustness checks on our results based on RDD, 2SLS

[3] 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.

and comparisons with repeat-sales.

4 Data

All the data used in this paper are publicly available, free of charge. Information on how to access the data is provided in Appendix A.

4.1 Matching EPC and PPD Datasets

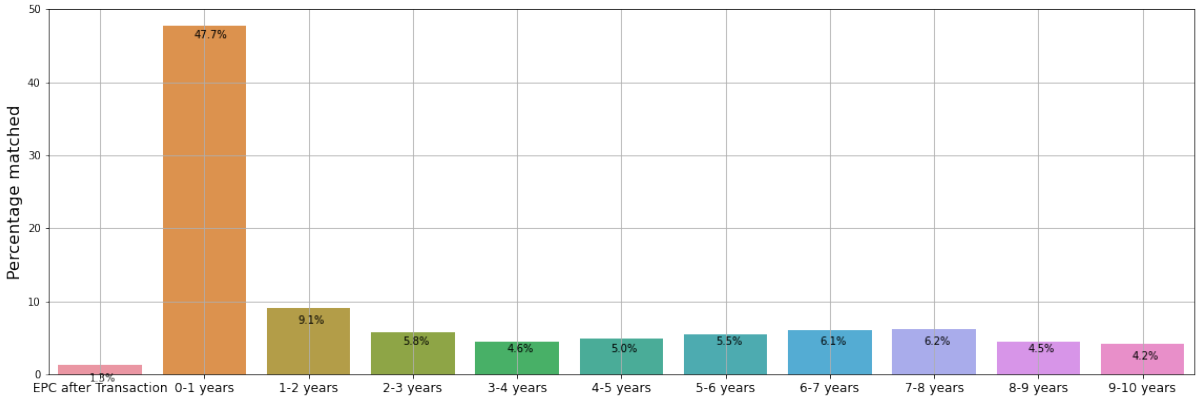
HM Land Registry Transaction Price Dataset

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, F = Flats/Maisonettes, O = Other), whether the property is a new build ("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 the Department for Levelling Up, Housing and Communities (2023b). Apart from a few exceptions, valid EPCs have been required when selling domestic properties since September 2008. 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: Figure 1 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 C1 in Appendix C 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 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).

Figure 2: Proportion of matched transactions

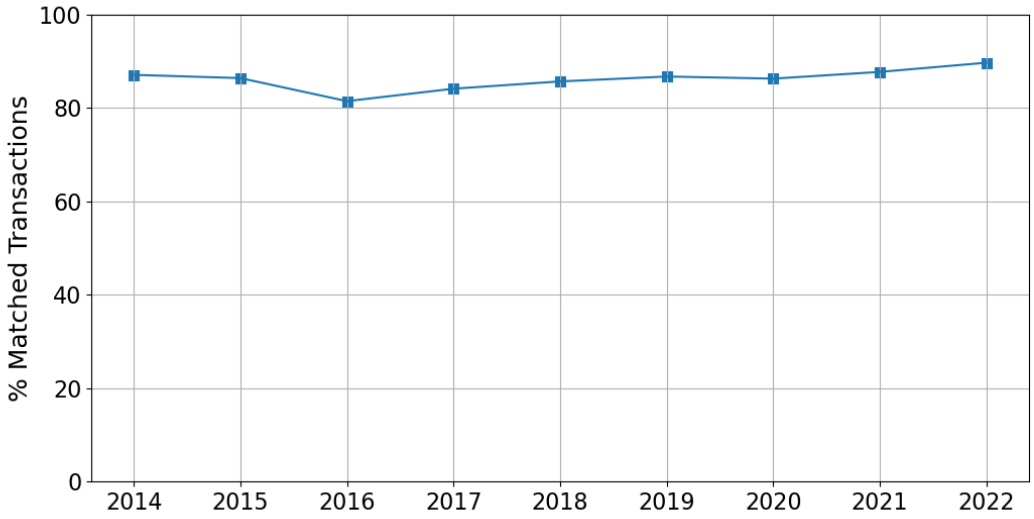


Figure 2 illustrates the percentage of transactions made for the years 2014 to 2022 that were matched to a corresponding EPC. Note that the transaction dataset includes commercial properties that are not included in the EPC sample.

EPCs are valid for ten years. 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.[4] 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.

The distribution of the energy efficiency scores of the matched data used to construct the hedonic model for each property type is compared with the distribution of the energy efficiency scores of the original EPC data in Figure B1 in Appendix B. As can be seen, for each property type, the distributions are similar, indicating that the matching process is not distorting the data.

4.2 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

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

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 (Ministry of Housing, Communities and Local Government, 2019). All 32 844 small areas or neighborhoods in England, called Lower-layer Super Output Areas are ranked. 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”. We separate the multiple deprivation index into 13 bands and then treat it as a dummy variable.

4.3 Data cleaning and summary statistics

Hedonic models can be sensitive to outliers (Rousseeuw and Leroy, 1987). We use Cook's distance to remove outliers in the dataset (Cook, 1979). This entails estimating hedonic models of log price separately for each property type (flat, semis/terraces, detached house) and region (All England, London, Rest of England), The exact list of characteristics included in the model is provided in section 5. Table 1 provides the summary statistics for some of the key characteristics in 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.

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)	254,009	257,878	428,301	418,778	272,336	268,500
Floor Area (mean)	63.4	63.1	127.6	126.3	90.8	90.2
Current Energy Efficiency (mean)	68.0	67.9	60.3	60.6	60.5	60.6
Potential Energy Efficiency (mean)	76.9	76.9	80.6	80.6	82.9	83.0
Built (share)						
Detached/Semi-Detached		38.1		100		55.9
Mid-Terrace		37.3		0		29.7
End-Terrace		24.6		0		14.4
Region (share)						
North East		2.9		3.1		5.0
North West		7.8		9.6		15.1
Y. and the H.		5.0		9.7		11.8
East Midlands		2.6		13.1		8.3
West Midlands		6.4		10.9		11.2
East		11.1		15.0		11.4
London		31.8		2.7		10.8
South East		21.5		21.4		16.2
South West		10.9		14.5		10.2
Observations	241,231	228,850	650,968	938,840	1,604,709	1,536,743

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	0.0	0.1	0.0
B	7.5	1.2	0.4
C	49.6	23.6	23.6
D	32.0	52.0	54.1
E	8.7	17.5	17.5
F	1.8	4.7	3.4
G	0.4	0.9	1.0

Note: Table 2 shows the share that belongs to each energy efficiency class using all market transactions between January 2015 and September 2023 that could be matched with an 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 is lower for newer properties.

Figure 3: Energy Efficiency By Age and Property Type

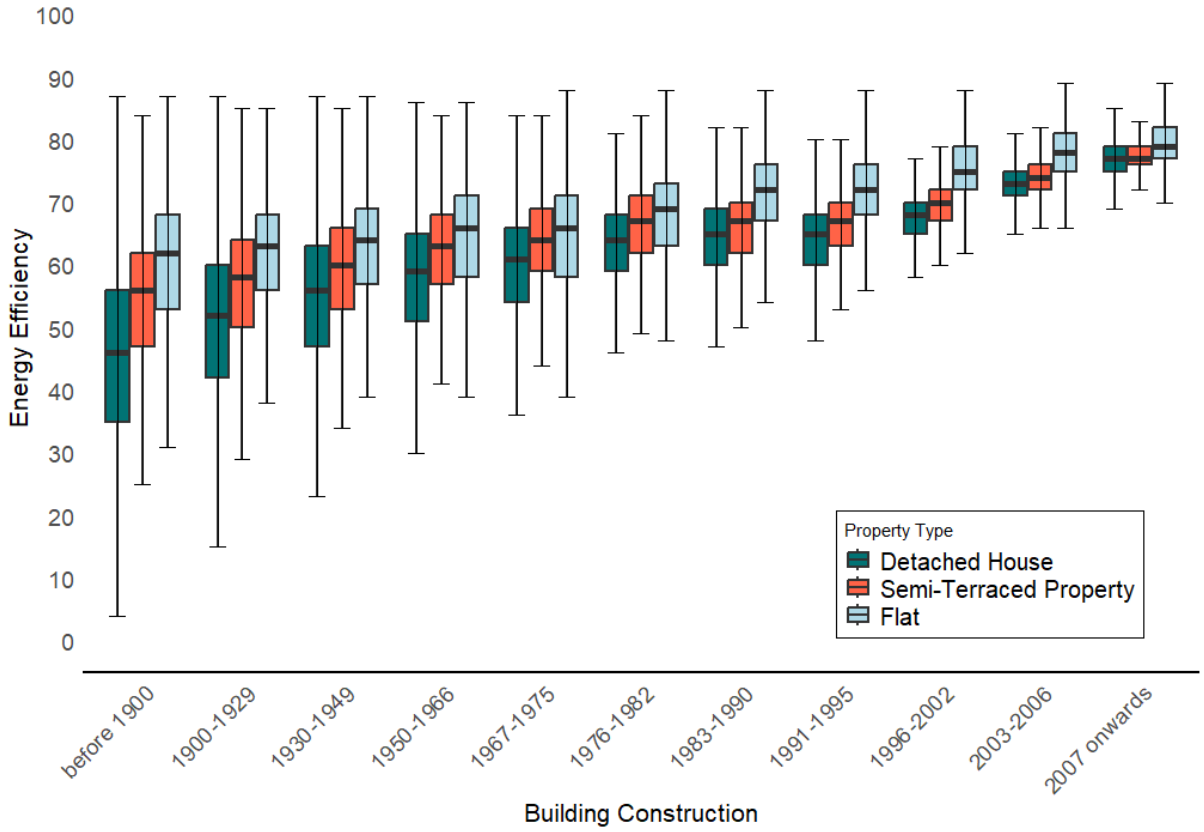


Figure 3 shows energy efficiency per property type for each building construction time range.

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 in England. 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
Flats	51.8	38.5	41.1	39.4	48.1	45.3	50.1	46.2	47.1
Semi/Terrace	67.0	72.8	73.2	66.8	72.4	65.3	78.5	63.3	63.1
Detached	46.1	61.0	63.2	62.1	66.5	65.8	81.1	68.2	67.8

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 2015-2023.

4.4 Recommended energy efficiency improvements

The list of all energy improvements included in the EPCs is provided in Appendix D in Table D1, while the 10 most recommended energy efficiency improvements for flats, semis/terraces and detached houses are shown in Table 4. 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 50mm 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: 10 Most Frequent Recommendations by Property Type

Recommendation – Flats	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
Recommendation – Semis/Terraces	**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
Recommendation – Detached Houses	**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 (2015-2023) that includes this particular recommendation. For example, 95.5% of EPCs recommend the installation of solar photovoltaic panels, 2.5kWp, for detached houses. The column "mean cost" estimates the average estimated cost over the same period (where the costs have not been uprated – see section 4.6).

4.5 Including an estimate of plot size for detached houses

Plot size is an essential determinant of the price of detached houses. However, neither the PPD- nor the EPC dataset includes any information about the size of the land the property is standing on. 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.[5]

Our proxy measure is based on the realisation that plot size is positively related to the distance between properties. The greater the distance to neighboring 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 neighbors. However, our hedonic regression results are quite robust to the exact number of UPRN neighbors used and whether we focus on distance or distance-squared as the variable of interest. Figure 4 graphically illustrates the concept behind our proxy plot variable.

Figure 4: Calculation of plot size approximation

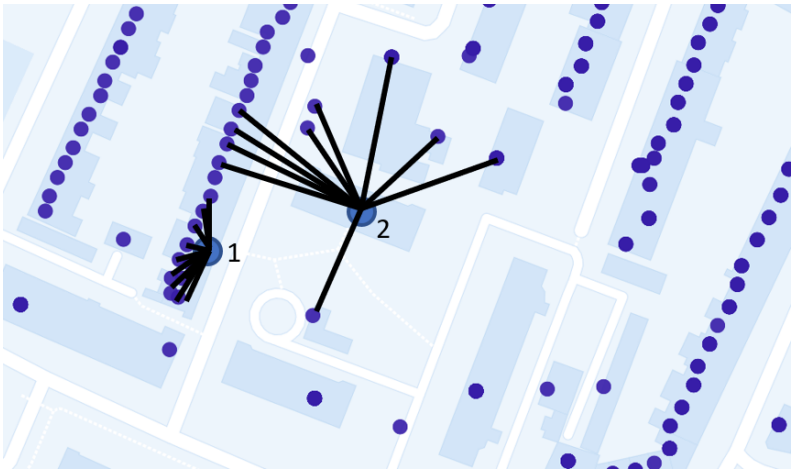


Figure 4 illustrates the calculation of the plot size proxy for detached houses. Each unique address is marked as a unique reference point (the small blue circles). Black solid lines depict the distances to the ten closest neighbors. Reference point 1 (left large blue circle) has a lower average distance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

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

Figure 5: Plot-Indicator Measurements for Detached Houses

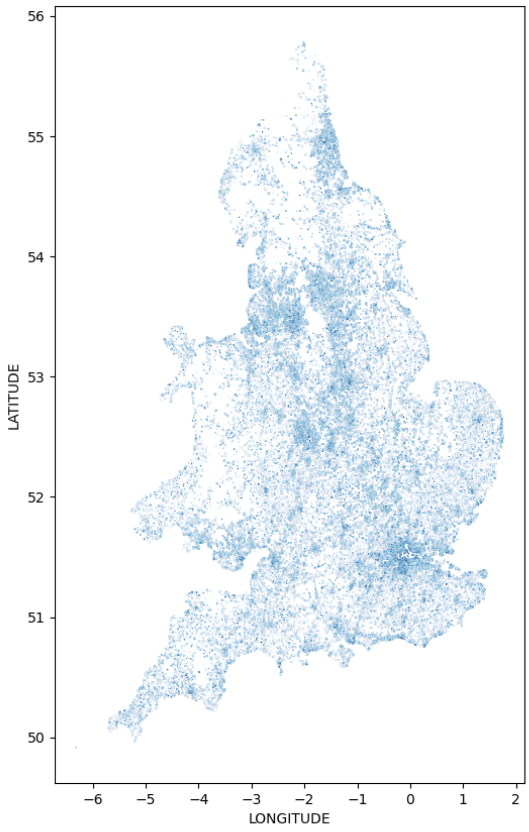


Figure 5 illustrates the plot-size indicator variable for detached houses in England and Wales. Darker dots indicate closer distance to neighbors.

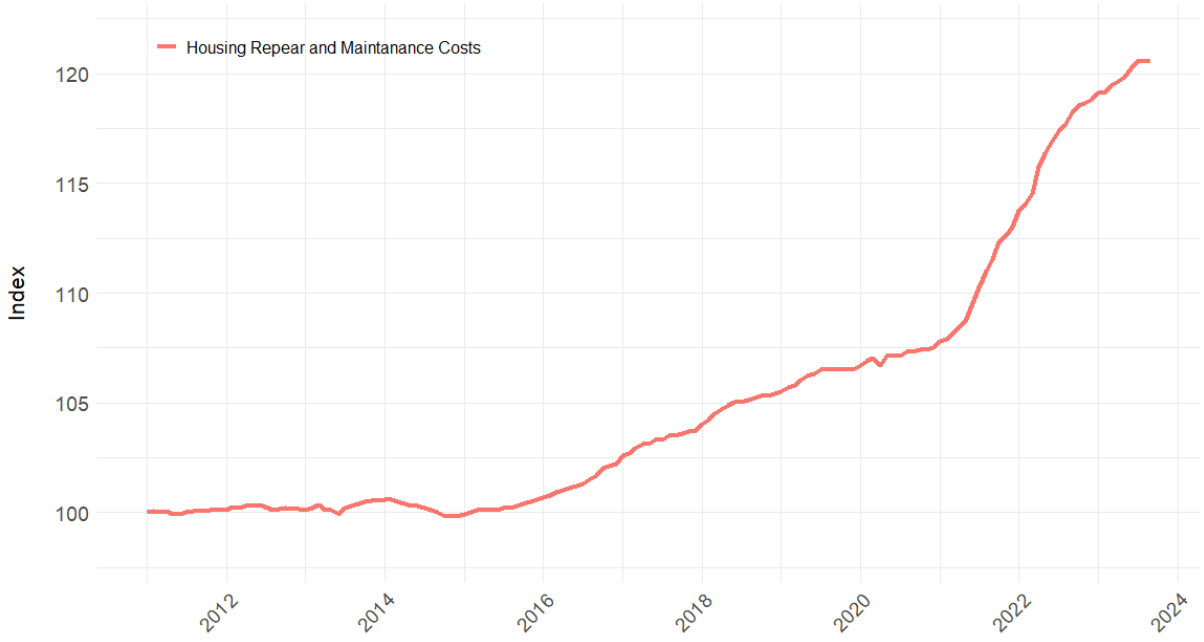
4.6 Uprating of costs

Starting in 2011, EPCs in England 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. EPC costs between 2011 and 2014 are uprated using the RPI: housing: repairs and maintenance charges index.[6] From 2014 onwards we use the construction price index: Housing repairs and maintenance.[7] Figure 6 represents this spliced cost index. The cost of housing repairs and maintenance increased by 25% between 2011 and 2023. By comparison, the official House Price Index produced by ONS rose by 44% over this same period. Without considering these cost increases, we would significantly overstate the market incentives for energy efficiency improvements in the later years of our sample.

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

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

Figure 6: Changes in the Cost of Housing Repair and Maintenance (2011-2023)



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.

5 Empirical Strategy and Results

5.1 A first look at the relationship between energy efficiency and price

An initial inspection of the data over the period 2015 to 2023 suggests a discrete jump as energy efficiency crosses the threshold between labels F and E (at the energy efficiency score of 39) and between E and D (at the energy efficiency score of 55). This can be seen in the left panel of Figure 7 which plots the average log price for properties at each energy efficiency level. A similar graph is shown in Sejas-Portillo et al. (2024). However, once all the prices are uprated to the same year-quarter, the observable jumps at the F-E and E-D boundaries disappear (as shown in the right panel of Figure 7).[8] This means that the jumps in the left panel are due to price observations just above the F-E and E-D thresholds tending to be from later periods while the observations just below the F-E and E-D thresholds tending to be from earlier periods. Once prices have been uprated, there is no clear evidence of a price jump at either the F-E or E-D thresholds.

[8] We use the official House Price Index of the Office of National Statistics (Office for National Statistics, 2024) to uprate (or downrate) all prices to April 2018.

Figure 7: Price and Energy Efficiency Score

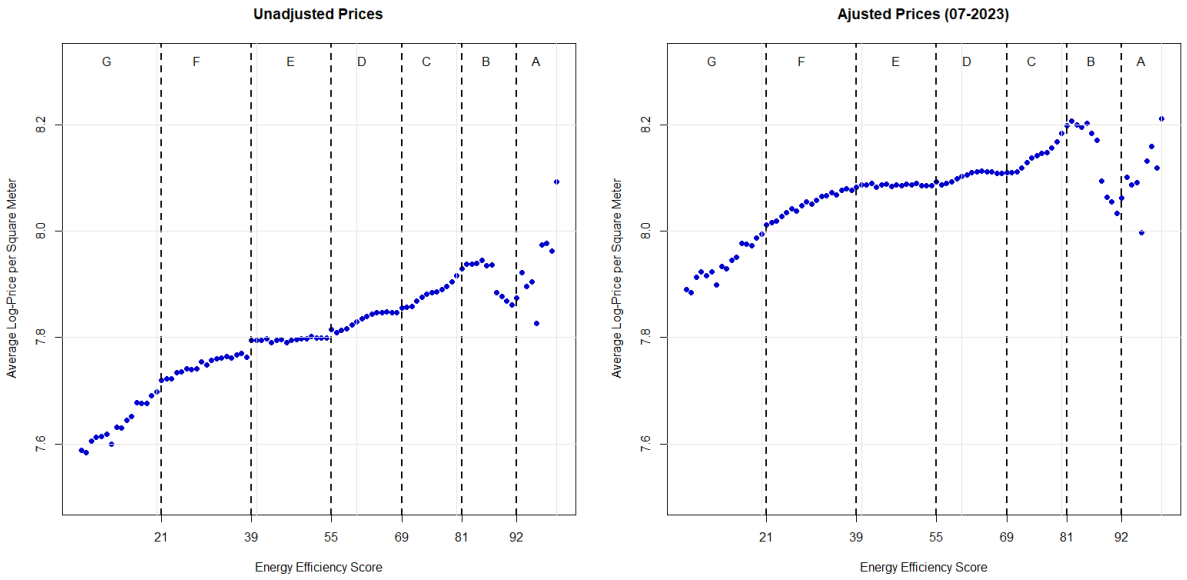


Figure 7 shows the relation between average log-prices and energy efficiency score calculated across all property types. The left panel are without adjustment of prices (whole time series), the right panel is with inflating the prices with the official ONS price index (Office for National Statistics, 2024).

The conclusion from this preliminary analysis is that the relationship between energy efficiency and log price (other things equal) is approximately linear and that there are no significant jumps at the label boundaries once time dummies are included (or house prices are uprated).

5.2 Hedonic regression

A hedonic model regresses the price (or log price) of a product (here a residential property) on a list of its characteristics. Hedonic regression can be used to estimate characteristic shadow prices of properties and to predict the impact on a property of a change in its characteristics. The conceptual basis of the hedonic approach was developed by Griliches (1961), Lancaster (1966) and Rosen (1974). In a housing context, we estimate separately for each region and property type (flat, semi/terrace or detached house) the following semi-log regression model on transacted properties over the period 2015-2023:

$$\ln(p_{t,n}) = \alpha + \theta E_{t,n} + \beta Q_{t,n} + \tau_t T_{t,n} + r_k R_{k,n} + \epsilon_{t,n}, \tag{7}$$

where $\ln(p_{t,k,n})$ is the log price of transaction n in period t, E_n is the property's energy efficiency level. The vector Q_n describes the observed characteristics of property n, $T_{t,n}$ denotes time dummies, $R_{k,n}$ denotes locational dummies and $\epsilon_{t,n}$ is an observation-specific error term with

mean zero. The full list of characteristics is as follows:
Energy efficiency score;
Floor area (in square meters);
Deprivation (sorted into 13 classes);
supergroup level (sorted into 8 classes);[9]
Local authority area;
Property type sub-level;
Construction year band;[10]
Tenure type (owner-occupied/private rental/social rental),
Leasehold/freehold;
Built form (detached, semi-detached, end-terrace, mid-terrace).

The main focus of interest here is the parameter θ, which measures the impact of the energy efficiency score on log price. Most of the existing 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 run from A to G). In papers that include the actual energy efficiency score from the EPC, there is some debate over 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. However, none of these papers use UK data. We use the linear specification for two reasons. First, our preliminary analysis from Figure 7 shows an approximately linear relationship between energy efficiency and price. Second, the estimated θ parameter in our hedonic model is reasonably stable when we interact it with the energy efficiency class, except for class G where the θ value is clearly lower (see Table E1 in Appendix E). 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 (7).

Once the hedonic model in (7) has been estimated, it can be used to predict the impact of energy efficiency improvements on 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 2015 to 2023. 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^{+}).

[9] "Rural", "Cosmopolitans", "Ethnicity", "Multicultural", "Urbanites", "Suburbanites", "Constrained", and "Hard-pressed".

[10] Before 1900, 1900-1929, 1930-1949, 1950-1966, 1967-1975, 1976-1982, 1983-1990, 1991-1995, 1996-2002, 2003-2006, 2007 onward.

These predicted prices are calculated as follows:

$$\hat{p}_{t,n} = \exp\left(\hat{\alpha} + \hat{\theta}E_n + \hat{\beta}Q_n^o + \hat{\tau}_t d_n + \hat{r}_k R_{t,k,n}\right), \tag{8}$$

$$\hat{p}_{t,n}^{+} = \exp\left(\hat{\alpha} + \hat{\theta}E_n^{+} + \hat{\beta}Q_n + \hat{\tau}_t d_n + \hat{r}_k R_{t,k,n}\right). \tag{9}$$

The expected price gain from implementing the recommended energy efficiency improvements listed in the property's EPC is given by $\hat{p}_{t,n}^{+} - \hat{p}_{t,n}$.

The hedonic regression results for flats are shown in Table 5. The estimated θ coefficient on energy efficiency is 0.0014 for All England, 0.0013 for London and 0.0013 for the Rest of England. These values imply that a 1% increase in energy efficiency raises property prices by between 0.13% and 0.14%. For semis/terraces the estimated θ coefficient on energy efficiency in Table 6 is 0.0026 for All England, 0.0012 for London and 0.0027 the Rest of England. For detached houses, the estimated θ coefficient in Table 7 is 0.0006 for All England, 0.0000 for London and 0.0011 the Rest of England. In all cases except detached houses in London, the estimated θ coefficients have p-values less than 0.001. These results suggest that prices are most responsive to energy efficiency improvements for semis/terraces and least responsive for detached houses. Also, prices are more responsive to energy efficiency improvements in the Rest of England than they are in London (except for flats).[11]

Table 5: Hedonic results for **Flats**

	All England	London	Rest of England
Energy efficiency score :	0.0014***	0.0013***	0.0013***
	[0.0001]	[0.0001]	[0.0001]
Dwelling characteristics	yes	yes	yes
Location characteristics	yes	yes	yes
Age	yes	yes	yes
Adj. R2	0.87	0.79	0.79
Obs.	226979	126359	154383

Note: Table 5 is constructed using transactions between 01.2015-07.2023. * p-value ≤ 0.05; ** p-value ≤ 0.01; *** p-value ≤ 0.001.

[11] More detailed summaries of the estimated parameters of the hedonic model are provided in Appendix F.

Table 6: Hedonic results for **Semis/Terraces**

	All England	London	Rest of England
Energy efficiency score :	0.0026***	0.0012***	0.0027***
	[0.0000]	[0.0001]	[0.0000]
Approx. neighborhood distance	1.0123***	2.1801***	1.0250***
	[0.0134]	[0.0753]	[0.0135]
Dwelling characteristics	yes	yes	yes
Location characteristics	yes	yes	yes
Age	yes	yes	yes
Adj. R2	0.90	0.88	0.86
Obs.	1519814	165840	1349149

Note: Table 6 is constructed using transactions between 01.2015-07.2023. * p-value ≤ 0.05; ** p-value ≤ 0.01; *** p-value ≤ 0.001.

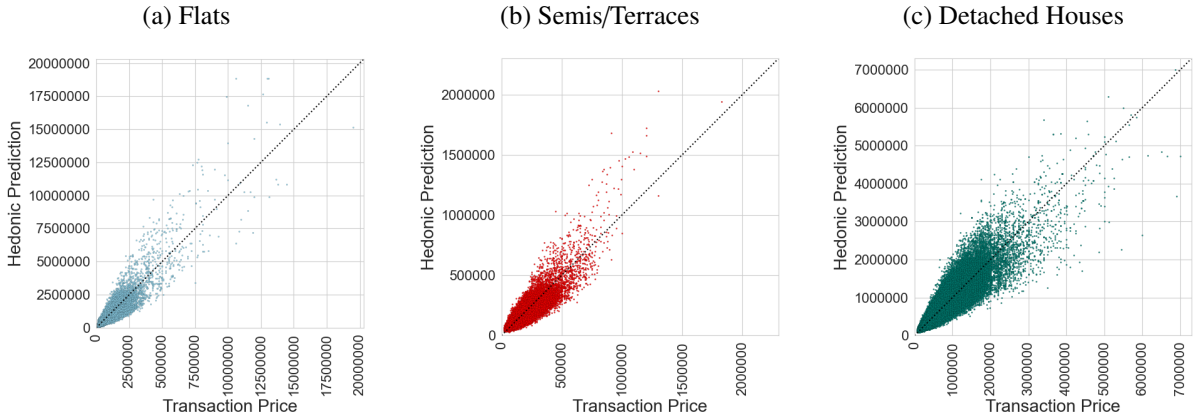
Table 7: Hedonic results for **Detached Houses**

	All England	London	Rest of England
Energy efficiency score :	0.0006***	0.0000	0.0011***
	[0.0000]	[0.0001]	[0.0000]
Approx. neighborhood distance	0.2481***	0.8580***	0.2458***
	[0.0065]	[0.0983]	[0.0065]
Dwelling characteristics	yes	yes	yes
Location characteristics	yes	yes	yes
Age	yes	yes	yes
Adj. R2	0.87	0.87	0.86
Obs.	634416	16550	617798

Note: Table 7 is constructed using transactions between 01.2015-07.2023. * p-value ≤ 0.05; ** p-value ≤ 0.01; *** p-value ≤ 0.001.

Figure 8 illustrates 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 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, semis/terraces and detached houses, respectively.

Figure 8: Prediction Performance of Hedonic Model (2015-2023)



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.

5.3 Does it pay to implement the EPC recommendations?

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 capitalization rate of energy efficiency improvements (CREEI) is defined as follows:

$$CREEI_{t,n} = 100\left[\frac{\left(\hat{p}^{+}_{t,n} - \hat{p}_{t,n}\right)}{k_{t,n}}\right]. \tag{10}$$

A CREEI below 100% implies that energy efficiency improvements are less than fully capitalized in property prices. For example, if the CREEI is 60%, this means that when all recommendations in the EPC are implemented, the price of the property increases on average at a rate of 60 Pounds for every 100 Pounds of expenditure on energy efficiency improvements.

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 ($\hat{p}_{t,n}$) rather than the actual price ($p_{t,n}$). 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 variables 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, 2013; Silver, 2018).

We derive aggregate measures of the CREEI by taking the median of the estimates in (10). The hedonic model in (7) is run separately for All England London and the Rest of England with data from 2015-2023. The implied median CREEIs for each property type calculated using (10) applied to the hedonic OLS model are shown in the top half of Table 8. For flats, the CREEIs are 59.2% for All England, 72.8% for London and 52.2% for the Rest of England. For example, the value of 59.2% implies that in All England on average 59.2% of the cost of increasing energy efficiency gets capitalized in the price of a property. This capitalization rate is higher in London and lower in the Rest of England. For semis/terraces the CREEIs are 62.4% for All England, 63.4% for London and 60.1% for the Rest of England. For detached houses the CREEIs are 21.4% for All England, 2.5% for London and 22.8% for the Rest of England.[12]

It is striking how much lower the CREEIs are for detached houses, especially in London. This may be due to the land leverage effect. The concept of land leverage refers to the share of the land in the total value of a property (Bostic et al., 2007). The land leverage is higher for detached houses than for semis/terraces and flats. It is especially high for detached houses in London. Given that energy efficiency improvements increase the value of the structure but do not affect the value of the land itself, it is not surprising that the CREEI is low for properties with especially high land leverage (e.g., detached houses in London).

Table 8: Capitalization Rate of Energy Efficiency Improvements (CREEI)

		England	London	Rest of England
Hedonic OLS	Flats	59.2%	72.8%	52.2%
	Semis/Terraces	62.4%	63.4%	60.1%
	Detached Houses	21.4%	2.5%	22.8%
2SLS	Flats	185.6%	206.8%	112.9%
	Semis/Terraces	148.8%	101.6%	152.6%
	Detached Houses	126.8%	234.9%	125.8%

Note: The formula for computing the CREEI is provided in (10). The average CREEI for flats in England obtained from the OLS hedonic model is $\boxed{59.2\%}$. This implies that when all recommendations in the EPC are implemented, the price of a flat rises on average at a rate of 59.2 Pounds for every 100 Pounds of expenditure on energy efficiency improvements.

[12] 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.

One weakness of the hedonic regression approach is that the numerator in the CREEI formula in (10) could be distorted by biases. First, the energy efficiency scores are measured subject to error (Jenkins et al., 2017; Hardy and Glew, 2019). Hardy and Glew (2019) find that between 37% and 62% of EPCs are inaccurate. This may cause downward bias in the estimated θ parameters (Hausman, 2001). Second, there may be omitted variables bias. The main concern here is that the unobserved quality of the structure and finishes could be positively correlated with energy efficiency, potentially imparting an upward bias to the estimated θ parameters. There may also be errors in the cost estimates of energy efficiency improvements (i.e., the denominator in (10).

In light of this discussion it is not clear in which direction the hedonic results and CREEIs may be biased. Next we consider regression discontinuity design (RDD), two-stage least squares and repeat-sales as robustness checks on our results.

5.4 Regression discontinuity design (RDD)

Regression Discontinuity Design (RDD) focuses on the price jumps as the energy efficiency of a property crosses a label threshold (e.g., from F to E). The assumption is that properties just below and just above a threshold should on average have similar levels of both the observed and omitted variables. Also, attenuation bias is avoided since the focus in the regression model is on the jump as a label threshold is crossed rather than on measuring the shadow price of a one unit increase in energy efficiency (Lee and Lemieux, 2010).

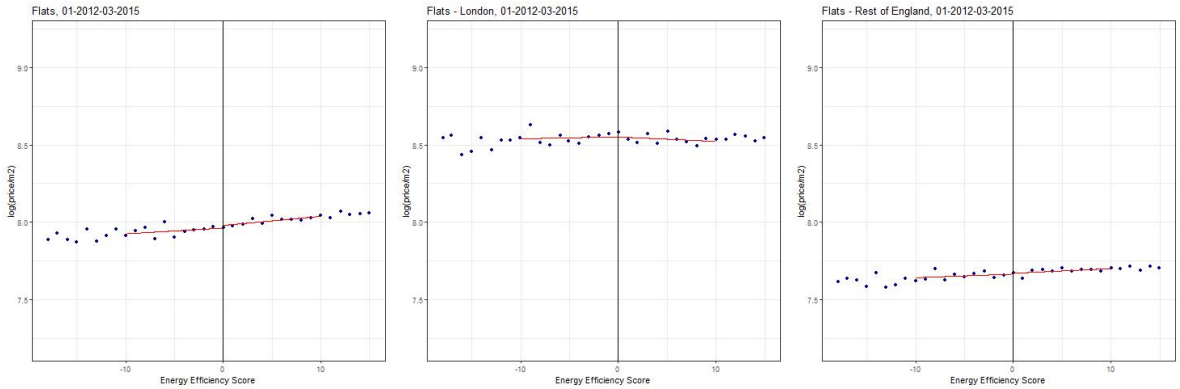
Unfortunately, our initial inspection of the data in Figure 7 did not find any clear evidence of price jumps at the label boundaries. However, a change in the law in 2018 provides an opportunity to investigate this issue further.

The UK government announced minimum energy efficiency standards (MEES) for rental properties in 2015 (HM Government, 2015). From 1st April 2018 all new rentals in England are required to achieve at least the E energy efficiency label. We focus our analysis in three ways. First, we concentrate on flats, since flats are more likely than other property types to be rented. Second, we focus specifically on the boundary between F and E referred to in the MEES. Third, we consider three time periods: pre-announcement (January 2012-March 2015), announcement (April 2015-March 2018), implementation (April 2018-December 2021). Graphs of relationship between energy efficiency and price in the vicinity of the F-E label boundary are shown in Figure 9 for All England, London and the Rest of England. The prices of all transactions are again uprated to April 2018 using the official House Price Index of the Office of National Statistics (Office for National Statistics, 2024).

Figure 9: RDD results of energy efficiency between energy score F and E for flats.

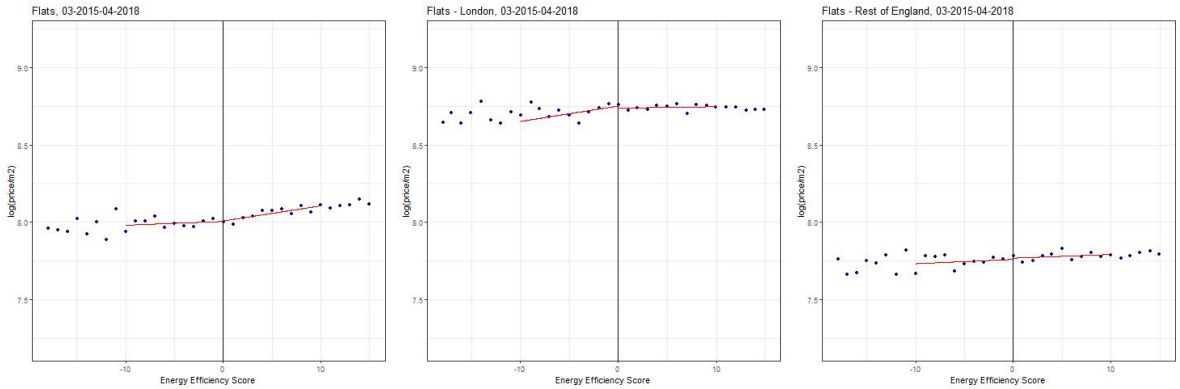
Pre-Announcement (Flats) - January 2012-March 2015

(a) All England (b) London (c) Rest of England



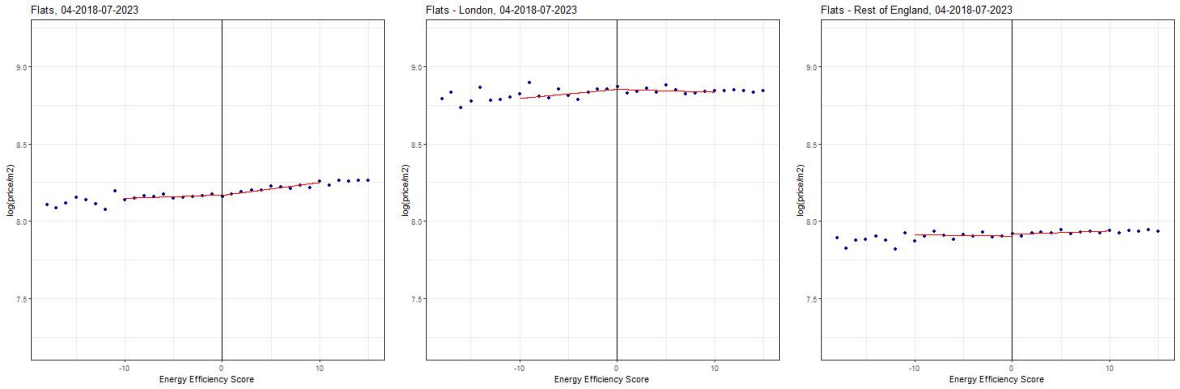
Announcement (Flats) - April 2015-March 2018

(d) All England (e) London (f) Rest of England



Implementation (Flats) - April 2018-December 2021

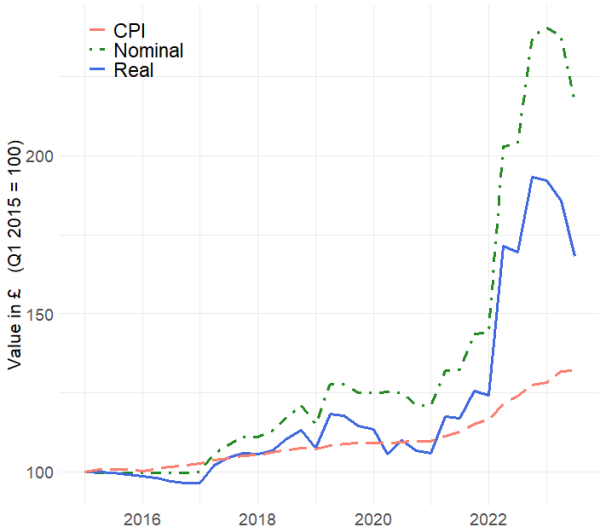
(g) All England (h) London (i) Rest of England



Note: No discernible jumps at the label boundary between F and E are observed once prices have been uprated.

There are no discernible price jumps at the F-E boundary in any of the pre-announcement, announcement or implementation phases in Figure 9. This finding is different from that of Sejas-Portillo et al. (2024), who also use UK data. This difference is presumably due to some combination of our use of uprated prices, different coverage time periods and our exclusion of Wales. We conclude that RDD is not applicable here. There is no evidence that buyers are influenced by the label itself over and above the actual energy efficiency score. In this sense, we

Figure 10: Electricity Price Development



do not observe any evidence of inattention. Market participants seem to respond to the actual energy efficiency score rather than changes in the label.[13]

5.5 Two-Stage Least Squares (2SLS)

Real energy prices rose by about 25% between 2015 and 2021. Then in 2022 they rose to nearly double the level of 2015 following Russia's invasion of Ukraine as shown in Figure 10. Here we use real energy prices as an instrument for energy efficiency in a 2SLS model. Given that we do not include time dummies in the 2SLS model, we first deflate all house prices using the official House Price Indices (HPIs) of the ONS. The HPIs used for uprating in the stage-2 equation are region and property-type specific.

We estimate the 2SLS model separately for each property type (i.e., flats, semis/terraces and houses) for All England, London and Rest of England over the period 2015-2023. In the first stage We also include interactions between the real energy price and energy efficiency labels.

$$E_{t,n} = \alpha_1 + \gamma C_c EP_{t-1} + Q_{t,n}\beta_1 + r_{1k}R_{t,n} + \epsilon_{t,n} \tag{11}$$

$$\ln(p^u_{t,n}) = \alpha_2 + \theta\hat{E}_{t,n} + Q^o_{t,n}\beta_2 + r_{2k}R_{t,n} + \varepsilon_{t,n} \tag{12}$$

In the first stage equation, $E_{t,n}$ is the energy efficiency score of property n sold in period t and EP_{t-1} is the real energy price one quarter earlier. We include a one quarter lag to allow energy efficiency to respond to changes in energy prices. C_t is a dummy variable for the energy

[13] See Myers (2019) and Sejas-Portillo et al. (2024) for a more detailed discussion of inattention.

efficiency label. The base label is F-G. The other categories are A-B, C, D and E. The other variables are the same as in (7). In the second stage equation, $\hat{E}$ is the predicted value of E obtained from the first stage and p^u is the uprated house price based on the official HPI of ONS for each property type/region (Office for National Statistics, 2024).

Table 9: 2SLS Results for **Flats** - Real Electricity Price as Instrument

	All England	London	Rest of England
Second stage			
Energy efficiency score :	0.0016***	0.0014***	0.0014***
	[0.0000]	[0.0001]	[0.0000]
Dwelling characteristics	yes	yes	yes
Location characteristics	yes	yes	yes
Age	yes	yes	yes
First stage			
Real Electricity Price$_{Q-1}$:	-0.3071***	-0.3147***	-0.3039***
	[0.0005]	[0.0010]	[0.0006]
A, B: Real Electricity Price$_{Q-1}$:	0.3861***	0.3962***	0.3818***
	[0.0005]	[0.0010]	[0.0006]
C: Real Electricity Price$_{Q-1}$:	0.3436***	0.3549***	0.3388***
	[0.0005]	[0.0009]	[0.0006]
D: Real Electricity Price$_{Q-1}$:	0.2699***	0.2825***	0.2644***
	[0.0005]	[0.0009]	[0.0006]
E: Real Electricity Price$_{Q-1}$:	0.1598***	0.1717***	0.1549***
	[0.0005]	[0.0009]	[0.0006]
F-stat for excluded instruments	281800	84550	197200
Number of observations	227949	72840	155110

Note: Table 9 is constructed using transactions between 01.2015-07.2023. * p-value ≤ 0.05; ** p-value ≤ 0.01; *** p-value ≤ 0.001.

Table 10: 2SLS Results for **Semis/Terraces** - Real Electricity Price as Instrument

	All England	London	Rest of England
Second stage			
Energy efficiency score :	0.0025***	0.0011***	0.0026***
	[0.0000]	[0.0000]	[0.0000]
Dwelling characteristics	yes	yes	yes
Location characteristics	yes	yes	yes
Age	yes	yes	yes
First stage			
Real Electricity Price$_{Q-1}$:	-0.2592***	-0.2480***	-0.2604***
	[0.0003]	[0.0006]	[0.0002]
A, B: Real Electricity Price$_{Q-1}$:	0.4183***	0.4060***	0.41865***
	[0.0007]	[0.0022]	[0.0004]
C: Real Electricity Price$_{Q-1}$:	0.3275***	0.3303***	0.32707***
	[0.0002]	[0.0004]	[0.0001]
D: Real Electricity Price$_{Q-1}$:	0.2689***	0.2688***	0.2688***
	[0.0002]	[0.0004]	[0.0001]
E: Real Electricity Price$_{Q-1}$:	0.1653***	0.1688***	0.16480***
	[0.0002]	[0.0004]	[0.0001]
F-stat for excluded instruments	182124	157000	1430000
Number of observations	1537931	166998	1370928

Note: Table 11 is constructed using transactions between 01.2015-07.2023. * p-value ≤ 0.05; ** p-value ≤ 0.01; *** p-value ≤ 0.001.

Table 11: 2SLS Results for **Detached Houses** - Real Electricity Price as Instrument

	All England	London	Rest of England
Second stage			
Energy efficiency score :	0.0006***	0.0001	0.0006***
	[0.0000]	[0.0001]	[0.0000]
Dwelling characteristics	yes	yes	yes
Location characteristics	yes	yes	yes
Age	yes	yes	yes
First stage			
Real Electricity Price$_{Q-1}$:	-0.2423***	-0.2000***	-0.2437***
	[0.0003]	[0.0019]	[0.0003]
A, B: Real Electricity Price$_{Q-1}$:	0.4044***	0.3886***	0.4051***
	[0.0004]	[0.0059]	[0.0004]
C: Real Electricity Price$_{Q-1}$:	0.3119***	0.3003***	0.3124***
	[0.0002]	[0.0014]	[0.0002]
D: Real Electricity Price$_{Q-1}$:	0.2547***	0.2369***	0.2553***
	[0.0002]	[0.0011]	[0.0002]
E: Real Electricity Price$_{Q-1}$:	0.1497***	0.1405*	0.1500***
	[0.0002]	[0.0011]	[0.0002]
F-stat for excluded instruments	748900	17190	73000
Number of observations	649707	16942	632765

Note: Table 11 is constructed using transactions between 01.2015-07.2023. * p-value ≤ 0.05; ** p-value ≤ 0.01; *** p-value ≤ 0.001.

The 2SLS results for flats, semis/terraces and detached houses are shown, respectively, in Table 9, Table 10 and Table 11. The Wu-Hausman test confirms the endogeneity of the energy efficiency scores. The high values of the F-test statistics for excluded instruments demonstrate the relevance of energy price as an instrument for energy efficiency (Stock and Yogo, 2005).

We undertake an informal test of the exclusion restriction, which following Kipperluis and Rietveld (2018) we refer to as the zero-first-stage test.[14] This test entails finding a subsample for which the effect of the instrument in the first stage is zero. The exclusion restriction then implies that the effect of the instrument on the outcome variable (property price) for this subsample should also be zero.

[14]This test has been used for example by Bound and Jaeger (2000), Altonji and Taber (2005) and Angrist and Schlosser (2010).

Here we use F and G properties as our subsample for the zero-first-stage test since these are properties for which energy efficiency is largely irresponsive to the energy price. In the first stage we re-estimate (11) but only on the F and G subsample of properties (and hence there are no interaction terms between the real energy price and energy efficiency labels). Therefore, the first-stage equation is as follows:

$$\ln(p_{t,n}^{\text{low}}) = \alpha_1 + \gamma EP_{t-1} + Q_{t,n}^{\text{low}}\beta_1 + r_{1k}R_{t,n}^{\text{low}} + \varepsilon_{t,n}^{\text{low}}, \tag{13}$$

where the superscript 'low' denotes that we are only considering properties with energy efficiency labels F and G. We are interested in cases where the estimated value of γ is not significantly different from zero. The results are shown in Table 12. It can be seen that the zero-first-stage condition is satisfied for flats and detached houses, but not by semis/terraces.

In the second stage we replace the predicted energy efficiency in (12) with the real energy price and again estimate the model only on the F and G subsample of properties as follows:

$$\ln(p_{t,n}) = \alpha_2 + \theta\hat{E}_{t,n} + \lambda EP_{t-1} + Q_{t,n}\beta_2 + r_{2k}R_{t,n} + \varepsilon_{t,n} \tag{14}$$

where λ captures the estimated direct effect of real energy prices on uprated property prices. In the second stage, for both flats and detached houses the energy price does not have a significant impact on property prices as required by the exclusion restriction.

Table 12: Zero-First-Stage Test

	Flats	Detached Houses	Semis/Terraces
Second stage			
Real Electricity Price	-0.0001	-0.0000	-0.0000
First stage			
Real Electricity Price	-0.0012	0.002	-0.0151***

Note: Table 12 is constructed using transactions between 01.2015-07.2023. * p-value ≤ 0.05; ** p-value ≤ 0.01; *** p-value ≤ 0.001.

Turning now to the 2SLS results, the estimated energy efficiency coefficients, θ, in stage 2 are almost identical to those obtained from the OLS hedonic models in Table 5), Table 6 and Table 7. This indicates that to the extent there are biases in the original hedonic results they broadly offset each other.[15] Hence the very low incentive for owners of detached houses in

[15] Attenuation bias and omitted variables bias may act in opposite directions. It should also be noted that OLS focuses on the Average Treatment Effect while 2SLS focuses on the Local Average Treatment Effect (Imbens and Angrist, 1994).

London to increase the energy efficiency of their properties still stands.

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

Repeat-sales samples can provide a further 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. 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 two transactions occur within 12 months as there is a higher likelihood that the property was purchased, renovated, and resold by a builder.[16] The repeat-sales sample we end up with contains 10 998 properties sold twice between 2015 and 2023.

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 (OLS or 2SLS) for periods t and $t+k$ for all properties in the two-EPC-repeat-sales sample. The average difference as a ratio is measured by K:

$$K = 100 \left(\prod_{n=1}^{N} \left(\frac{p_{t+k,n}(E_{t+k,n})}{p_{t,n}(E_{t,n})} \Big/ \frac{\hat{p}_{t+k,n}(E_{t+k,n})}{\hat{p}_{t,n}(E_{t,n})} \right)^{1/N} \right). \tag{15}$$

A value of K greater than 100 in Table 13 means that on average the actual repeat-sales price increase is larger than the predicted increase obtained from the OLS or 2SLS hedonic models. The repeat-sales price increases are between 9% and 16% higher depending on the property type.

[16] Omitting higher frequency repeat sales is a standard approach when constructing repeat-sales price indices (S&P Dow Jones Indices, 2022).

Table 13: Comparing Hedonic OLS and Instrumental Variable (IV) with Repeat-Sales Price Changes

Property type	OLS	2SLS	Repeat Sales Observations
Flats	109.8		996
Semis/Terracess	116.1		7997
Detached Houses	112.7		2005

Note: "Sample" denotes the number of properties over which each coefficient is calculated. The samples cover transactions over the period 2015-2023. The repeat sales results show a higher return on energy efficiency improvements than under both the OLS and 2SLS hedonic approaches.

The repeat-sales results provide an independent check on the hedonic results, although only for a small subsample of the full dataset. Also, actual energy efficiency improvements often occur as part of broader renovations. While we restrict our repeat-sales comparison to properties whose observable characteristics do not change, this will not completely remove the effect of renovations. Hence the observed repeat-sales price increases should be viewed as an upper bound on the actual impact of energy efficiency improvements on property prices. With these provisos in mind, the results from the repeat-sales robustness check look quite encouraging. The repeat-sales based price increases from energy efficiency improvements are on average 10% larger for flats, 16% larger for semis/terraces and 13% larger for detached houses as compared with the results obtained from the hedonic approach.

6 Conclusion

Matching properties across Energy Performance Certificates (EPCs) and the HM Land Registry in England 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 stated in the EPCs, we were able to measure the Capitalization 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. Therefore, a proxy measure of plot size is needed to obtain meaningful results for detached houses. We constructed this proxy plot size variable based on each property's distances to its ten nearest neighbors. 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 then calculated CREEIs using OLS and 2SLS hedonic approaches. The OLS and 2SLS results were very similar. The CREEI is 23% for detached houses, 52% for flats and 60% for semis/terraces. This means that between 23% and 60% of the expenditure on energy efficiency improvements is capitalized in property prices. As a robustness check we compared the predicted price increases of energy efficiency improvements obtained from our hedonic models with actual price increases from repeat sales. The actual repeat-sales price increases were between 9% and 16% higher depending on the property type. Given repeat-sales price increases associated with energy efficiency improvements may also include other renovations, these estimates should be treated as an upper bound. In this sense, the hedonic results look quite plausible.

One issue for future research is to investigate more closely the case of detached houses in London where the capitalization rate of energy efficiency improvements was only 2%. This may be due to the land share of the total value of a property being high in London (i.e., high land leverage) (Bostic et al., 2007) or a confounding effect of interactions between unobserved quality and location. A second issue is that we focus on the case where owners implement all the recommendations in the EPCs. Some recommendations are likely to yield higher CREEIs than others. It would be interesting therefore to revisit this topic and determine which individual improvements generate the highest CREEIs.

References

Altonji, J. G., T. E. E. and Taber, C. R. (2005). An evaluation of instrumental variable strategies for estimating the effects of catholic schooling. Journal of Human Resources, 40:791—-821.

Angrist, J. D., V. L. and Schlosser, A. (2010). Multiple experiments for the causal link between the quantity and quality of children. Journal of Labor Economics, 28:773—-824.

Asproudis, E., Gedikli, C., Talavera, O., and Yilmaz, O. (2024). Returns to solar panels in the housing market: A meta learner approach. Energy Economics, 137:107768.

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.

Bound, J. and Jaeger, D. A. (2000). Do compulsory school attendance laws alone explain the association between quarter of birth and earnings? In Research in Labor Economics, Volume 19, 83–108. Bingley, West-Yorkshire: Emerald Group Publishing Limited.

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 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 (2023a). Energy performance of buildings data - register. January 20, 2023, https://epc.opendatacommunities.org/.

Department for Levelling Up, Housing and Communities (2023b). 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. In Diewert, W. E., Greenlees, J. S., and Hulten, C. R., editors, Price Index Concepts and Measurement. University of Chicago Press, pages 445–500.

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 (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.

Griliches, Z. (1961). Hedonic price indexes for automobiles: an econometric analysis of quality change. In The Price Statistics of the Federal Government, General Series No. 73. New York: NBER.

Hardy, A. and Glew, D. (2019). An analysis of errors in the energy performance certificate database. Energy Policy, 129:1168–1178.

Hausman, J. A. (2001). Mismeasured variables in econometric analysis: problems from the right and problems from the left. Journal of Economic Perspectives, 15(4):57–67.

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

HM Government (2015). The energy efficiency (private rented property) (england and wales) regulations 2015.

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.

Imbens, G. W. and Angrist, J. D. (1994). Identification and estimation of local average treatment effects. Econometrica, 62:467–475.

Jenkins, D., Simpson, S., and Peacock, A. (2017). Investigating the consistency and quality of epc ratings and assessments. Energy, 138:480–489.

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.

Kipperluis, H. v. and Rietveld, C. A. (2018). Beyond plausibly exogenous. Econometrics Journal, 21:316–331.

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.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. Journal of Economic Literature, 48(2):281–355.

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.

Myers, E. (2019). Are home buyers inattentive? evidence from capitalization of energy costs. American Economic Journal: Economic Policy, 11(2):165–188.

Myers, E., Puller, S. L., and West, J. (2022). Mandatory energy efficiency disclosure in housing markets. American Economic Journal: Economic Policy, 14(4):453–487.

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.

Office for National Statistics (2024). Uk house price index: reports. 17 July, 2024. https://www.gov.uk/government/collections/uk-house-price-index-reports.

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.

Sejas-Portillo, R., Moro, M., and Stowasser, T. (2024). The simpler the better? threshold effects of energy labels on property prices and energy-efficiency investments. American Economic Journal: Applied Economics, forthcoming.

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.

Stock, J. H. and Yogo, M. (2005). Testing for weak instruments in linear iv regression. In Andrews, D. W. K. and Stock, J. H., editors, Identification and Inference For Econometric Models: Essays in Honor of Thomas Rothenberg. Cambridge, UK: Cambridge University Press.

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.

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: How to Access the Data Used in this Paper

All the data used in this paper are publicly available, free of charge. Guidelines for accessing the data are provided below.

(i) The Energy Performance Certificate (EPC) dataset

Information and instructions for accessing the EPC data are available from the following website: Department for Levelling Up, Housing and Communities (2023b). Users are required to register to access the data (Department for Levelling Up, Housing and Communities, 2023a).

(ii) The Price Paid Dataset (PPD)

Information about the PPD is available from the following website: HM Land Registry (2023). The PPD can be downloaded directly from this website in either csv or txt format.

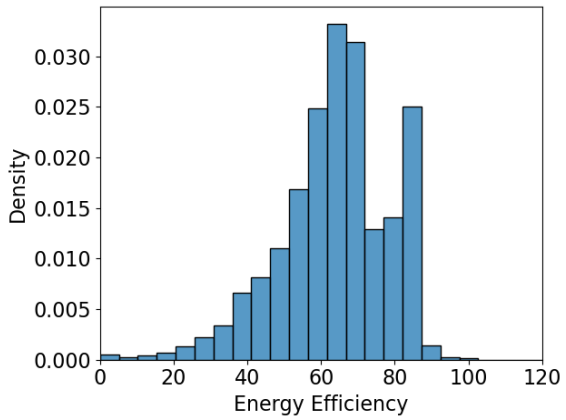
(iii) Area Classifications

We use two area classifications in addition to the Local Authority Districts provided in the EPCs and PPD. The first consists of deprivation indices for England, which can be downloaded from the following website: Ministry of Housing, Communities and Local Government (2019). The second consists of supergroups corresponding to the following categories: "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". The supergroup data can be downloaded from the following website: Office for National Statistics (2023).

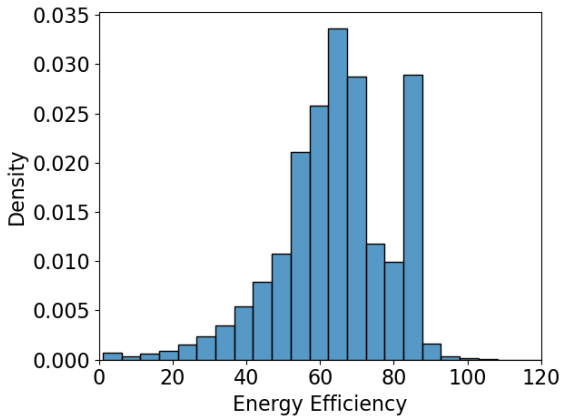
Appendix B: Distribution of Energy Efficiency

Figure B1: Energy Efficiency: merged transaction data (left), merged EPC data (right)

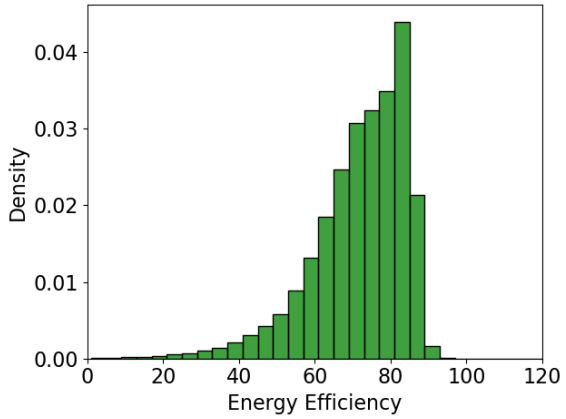
(a) Energy Efficiency for Detached Houses: 1 510 885 observations



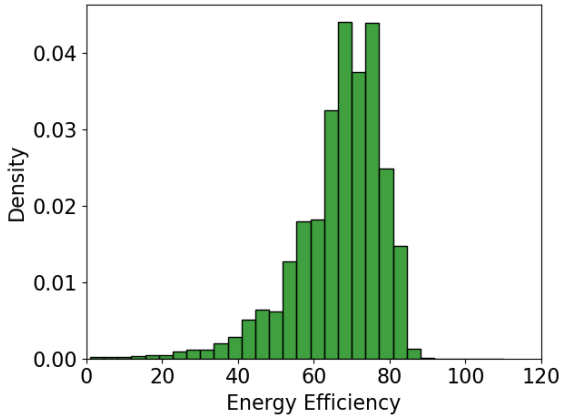
(b) Energy Efficiency for Detached Houses: 1 335 157 observations



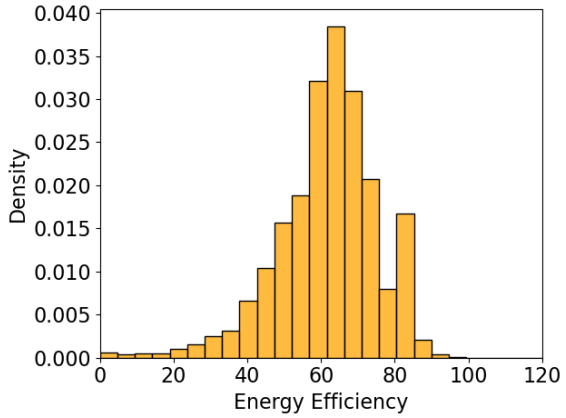
(c) Energy Efficiency for Flats: 730 579 observations



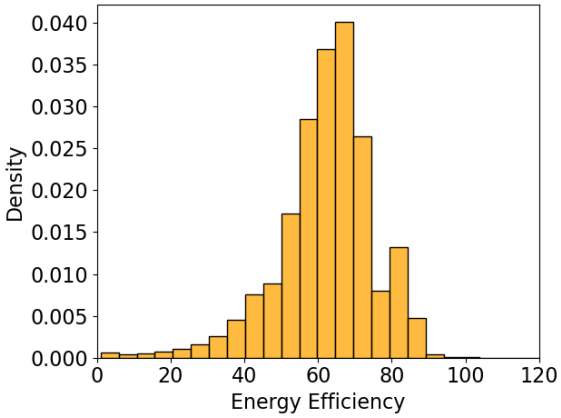
(d) Energy Efficiency for Flats: 559 200 observations



(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 C: Matching Rates

Table C1: 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 D: Possible Energy Efficiency Improvements Included in EPCs

Table D1: 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

Appendix E: Non-Linearity Check of Energy Efficiency Coefficient

Table E1: Regression results for interaction term of energy efficiency score and rating class

Energy rating	Flats	Semis/Terraces	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 E1 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 F: Detailed Results for the Hedonic Models