Energy performance ratings and house prices in Wales: an empirical study

Article

Accepted Version

ENERGY PERFORMANCE RATINGS AND HOUSE PRICES IN WALES: AN EMPIRICAL STUDY

Dr Franz Fuerst

Department of Land Economy

University of Cambridge

ff274@cam.ac.uk

Prof Pat McAllister

School of Real Estate & Planning

University of Reading

p.m.mcallister@reading.ac.uk

Dr Anupam Nanda

School of Real Estate & Planning

University of Reading

a.nanda@reading.ac.uk

Prof Pete Wyatt

School of Real Estate & Planning

University of Reading

p.wyatt@reading.ac.uk

Abstract

This paper investigates the price effect of Energy Performance Certificate (EPC) ratings on residential dwelling prices in Wales. Drawing on a sample of approximately 192,000 transactions, the capitalisation of energy efficiency ratings into house prices is investigated using two approaches. The first adopts a cross-sectional framework to investigate the effect of EPC rating on price. The second approach is based on a repeat-sales methodology to investigate the impact of EPC rating on house price appreciation. Statistically significant positive price premiums are estimated for dwellings in EPC bands A/B (12.8%) and C (3.5%) compared to houses in band D. For dwellings in band E (-3.6%) and F (-6.5%) there are statistically significant discounts. Such effects may not be the result of energy performance alone. In addition to energy cost differences, the price effect may be due to other benefits of energy efficient features. An analysis of the private rental segment reveals that, in contrast to the general market, low-EPC rated properties were not traded at a significant discount. This suggests different implicit prices of potential energy savings for landlords and owner-occupiers.

Key Words

energy; performance; certificates; dwellings; prices; Wales

Acknowledgements

We are grateful to the Welsh Low/Zero Carbon Hub for providing the financial support for this research.

1 Introduction

In most developed economies environmental certification schemes have been introduced in the commercial and the residential property sectors. These schemes are a market-based mechanism designed to inform consumers about the environmental performance of a product. This information is then expected to influence consumer behaviour, increase demand for less environmentally harmful products, produce changes in the relative supply of energy efficient products and, ultimately, reduce environmental impacts. Market prices are important in that they send demand signals from consumers to suppliers about what, where and when to produce. In particular, price premiums provide an economic incentive for producers to innovate and incur any additional production costs associated with improved energy performance. A key issue is the extent to which, within the purchase decision and associated price determination, consumers are willing to pay a premium for good environmental performance. The focus of this paper is on the price effects of energy performance in the residential property sector. If a price premium can be attributed to energy efficiency in the housing market, then, depending on the trade-off with additional costs, it may provide residential developers with evidence to justify the supply of more energy efficient dwellings and incentives for existing owners to improve the environmental performance of their homes and investments.

In 2008 the measurement of energy use in new and existing buildings in the UK became obligatory following the implementation of the European Union's Energy Performance of Buildings Directive. This required all buildings at the point of construction completion, sale or rent (or every ten years) to be issued with certificates that provide information about their energy performance. These Energy Performance Certificates or EPCs are asset ratings intended to inform potential purchasers about the intrinsic energy performance of a building and its associated services. The residential property market is by far the richest real estate sector in terms of transaction volume and, with seven years of recorded EPCs, there is sufficient scope to introduce a variety of statistical methods to control for price determinants other than energy performance. Using Wales as a case study area for the first time, in this paper we use a large sample of relatively homogeneous residential dwellings to investigate whether EPCs influence transaction prices and price growth rates. Because the data set included repeat sales we were able to exercise a greater degree of control for potential bias from dwelling-specific fixed effects. Furthermore, a distinction was drawn between purchasers who acquire dwellings for their own occupation and those who acquire for investment reasons – 'buy-to-let' landlords who lease dwellings to tenants – in order to investigate whether there was a significant difference in energy efficiency price premium between the two groups. This distinction is important because in 2013 buy-to-let landlords owned 19% of all dwellings in the UK compared to 11% a decade earlier[1]. This growing category of investors may value energy efficiency differently as, under typical lease arrangements, tenants usually pay energy bills. The empirical research includes a series of robustness checks to try and control for potential omitted variable bias due to properties that may have been improved or may be in very good or very poor condition.

The remainder of the paper is structured as follows. The next section reviews previous studies of the price impact of environmental performance labels on residential dwelling prices. Section 3 describes the data set and modelling approach used in this study. Essentially it examines the capitalisation of energy efficiency ratings into house prices using two approaches. The first adopts a cross-sectional framework to investigate the effect of EPC band (and EPC rating) on a large sample of dwelling transactions. The second approach is based on a repeat-sales methodology to examine the impact of EPC band and rating on house price appreciation. Section 4 presents the results and section 5 provides some discussion of the findings before concluding comments in the final section.

[1] Department for Communities and Local Government (2015) Table 101: Dwelling stock by tenure, https://www.gov.uk/government/statistical-data-sets/live-tables-on-dwelling-stock-including-vacants, accessed 22 October 2015

2 Literature Review

Following the energy crises of the 1970s, some of the earliest relevant literature investigated the relationship between energy efficiency and residential prices (see Laquatra *et al*, 2002 for a review). A body of US work from the 1980s broadly identified a positive relationship between energy efficiency and residential sale prices (see Halvorsen and Pollakowski, 1981; Johnson and Kaserman, 1983; Quigley, 1984; Laquatra, 1986; Dinan and Miranowski, 1989; Quigley and Rubinfeld, 1989). However, in the last decade, growing concern about climate change has stimulated another wave of research on energy performance and residential sale prices. Given the rapid growth of research in this area, below we review the work most closely related to this study.

In a largely overlooked initial study, the Australian Bureau of Statistics (2008) examined residential sales in the Australian Capital Territory for the period 2005 (2,385 transactions) and 2006 (2,719 transactions). For the 2005 sample, it found an approximately 1% price premium for every 0.5 increase in the Energy Efficiency Rating (EER), which ranges from 0 to 5. For the 2006 sample, there was an approximately 2% premium for every 0.5 increase in EER. For the pooled sample, relative to a zero rating, premiums of 1.6% (EER 1), 3% (EER 2), 5.9% (EER 3), 6.3% ((EER 4) and 6.1% (EER 5) were found; the marginal addition to the premium declining as rating increased. The explanatory power of model was high and there was a large range of controls for the quality of the dwellings.

Kahn and Kok (2014) conducted a hedonic pricing analysis of all single-family home sales in California between 2007 and 2012. Using a sample of matched properties based on the likelihood of having a green label and the local area weather condition, they found a 2% premium for green labels. While the perennial difficulty of measuring unobserved non-financial benefits of green label still remains, this study shows a robust positive association based on several alternative specifications. However, the results are based on comparing a relatively small ’treated’ sample with a substantially larger ‘non-treated’ sample.

With an interesting focus on presale (dwellings bought from developers) and resale (dwellings sold by owners) prices, Deng and Wu (2014) compared a sample of 13,224 dwellings in 62 Green Mark developments with 55,983 dwellings in 1,375 non-Green Mark developments in Singapore between 2000 and 2010. They applied a range of approaches including hedonic methods (supplemented by PSM) and difference-in-difference (DID) methods to investigate the price effects of the Green Mark certification. Similar to Deng *et al.* (2012), overall they estimate an average price premium of about 4-5%. In terms of the different levels of award, the estimated premium for the Platinum rating was 11%; the comparable figures for Gold and Certified ratings were 5% and 1.6% respectively. Premiums for resales found to be substantially higher. Using a smaller sample of repeat transactions, DID approach estimates price appreciation premium for Green Mark dwellings of 2-3%. They infer from the results that developers are capturing a small part of the green premium. However, without details of costs of achieving certification, similar to most previous studies, they were unable to assess whether the price premium compensated developers for additional costs.

In Europe, based on a sample of 31,993 residential sale prices in the Netherlands in 2008-9 for dwellings with (voluntary) EPC ratings, Brounen and Kok (2011) identified premiums of 10%, 5.5% and 2.5% for A, B and C respectively, compared to D-rated dwellings. For dwellings rated E, F and G, there were respective discounts of 0.5%, 2.5% and 5%. The data set contained a broad range of control variables including dwelling size, insulation quality, central heating and level of maintenance. Using a composite sustainability metric based on 36 variables to provide a sustainability score for each dwelling, Feige *et al.* (2013) drew upon rental prices of a sample of 2,453 residential apartments in Switzerland. Their results revealed that some sustainability-related features had significantly positive effects, others had no effect on price and some had a negative effect. Importantly in the context of this paper they found an unexpected negative relationship between energy efficiency and price. This was attributed to Swiss residential lease structures where landlords tend to recover the estimated cost of energy from tenants in advance. Hence, less energy efficient buildings may have appeared to have a higher rent since the energy cost is 'bundled' with rent. Kholodilin and Michelsen

(2014) investigated the residential rental market in Berlin and found that energy efficiency savings are generally capitalised into prices and rents and that buyers are able to anticipate energy and house price movements. Another finding relevant to this paper is the significantly lower implicit prices of energy efficiency of rental dwellings compared to owner-occupied dwellings. The authors explain this difference as a sign of the market power of tenants or as a result of the split incentive problem. Similarly, Cajias and Piazolo (2013) find higher total returns and higher rents for energy-efficient dwellings in their study of the German housing market between 2008 and 2010. They estimate that a one percent energy saving raises rents by 0.08% and the market value of a property by 0.45%. Hyland *et al* (2013) analysed the impact of energy efficiency ratings in Ireland on residential asking prices and rental rates based on a rich data set of Building Energy Ratings (the Irish equivalent of the EPC) as well as property and price information. They found asking price premiums relative to D-rated properties for A (9%), B (5%) and C (1.7%). There was no significant discount for E-rated dwellings and a discount of approximately 11% for F/G. Rental premiums were 1.8% for A and B rated dwellings compared to D and no significant price effect on C-rated dwellings. There were rental discounts for E (1.9%) and F/G (3.2%) rated dwellings. The analysis does not appear to control for age of buildings and as a result there may be a risk of misattributing age effects to energy efficiency effects.

The European Commission (2013) published a report that included a series of studies of the effect of EPC rating on prices in a range of European countries. Focusing on the residential sector, in Austria, using a sample of 1,077 rents and 2,246 sale prices mainly in Greater Vienna, an approximately 8% increase in price and 4% increase in rent was found for each change in EPC rating. In Belgium, based on 26,000 listed prices, the majority of which were in Flanders, a 3-5% increase in price was associated with an increase in every 100 CPEB (Certificat de Performance Energétique des Bâtiments) points. In France, samples of 1,263 and 1,915 sale transactions for Marseilles and Lille respectively revealed an approximately 4% increase in price per change in EPC rating for sale prices. In Ireland approximately 26,500 listed rental prices and 11,000 listed sale prices revealed an approximately 2.8% increase in price per change in EPC rating for selling prices and 1.4% increase in price per change in EPC rating for rental prices.

Finally, in a study closely related to this paper, drawing upon a sample of 325,950 English housing transactions with mandatory energy certificates and with a control for age of dwelling, Fuerst *et al* (2015) significant positive price premiums for dwellings with EPC ratings of A/B (5%) or C (1.8%) compared to dwellings rated D. For dwellings rated E and F statistically significant discounts were estimated, -0.7% and -0.9% respectively. Dwellings rated G sold for approximately 6% less. Turning to price growth, the findings were less clear-cut. Dwellings in EPC band C experienced significantly higher price *growth* than those in band D. However, this was not the case for the dwellings in bands A and B, which experienced significant price depreciation compared to D-rated dwellings. Dwellings in band E (-0.18%) and F (-0.26%) were also estimated to have had statistically significant lower rates of price growth compared to D-rated dwellings. Similar to previous studies, the study did not control for potential omitted variable bias due to the absence of information on improvements and other quality variables such as age and condition of bathrooms and kitchens.

The private rented sector of the residential housing market is of particular interest in that energy costs are not borne by the buyer but by the tenant. Buy-to-let landlords may value energy efficiency only to the extent that it enables them to charge higher rents, achieve shorter vacancy periods or otherwise increase the attractiveness of their investment. Due to the well-documented split incentive problem, it has been observed in past studies that rental properties generally achieve a lower energy efficiency standard than owner-occupied properties (Rehdanz 2007). This may due, at least in part, to the general quality of the rental stock being poorer. For example, Iwata and Yamaga (2008) purport that the optimal condition of a rented dwelling is lower if an investor expects heavy utilisation of the dwelling by the tenant.

Some general points about these studies can be made. First, hedonic model estimates can be sensitive to choice of model specification and availability of information on variables that determine prices. This is particularly so if it is suspected that the price impact of an attribute (energy performance, for example) is likely to be small in comparison to other attributes such as location and age of dwelling. Data availability is, therefore, a major challenge to researchers in this area. The feasibility and quality of empirical research into the price effects of energy efficiency certification is dependent upon the availability of dwelling-level data on three main areas: market prices (rents and sales), environmental performance of real estate assets, and building attributes. Data constraints may mean that certain attributes are omitted from the hedonic model and this can lead to bias. For instance if age of dwelling is omitted and age and energy performance are considered to be correlated, negative price effects associated with aging would be reflected in the energy efficiency variable. Alternatively, being energy efficient may only be one part of a bundle of 'extras' that a housing developer has used to create a superior product. For instance, homes with better energy performance may be of a higher quality of construction. By omitting this variable (superior construction) a construction quality price effect could be misattributed as an energy efficiency price effect.

Second, house purchasers can obtain a bundle of costs and benefits when they buy energy efficient homes that are not energy related. Certain attributes that enhance energy or environmental performance can also enhance other aspects of performance. For instance, houses need to have double-glazing and/or a modern water heating system in order to obtain a good EPC rating. However, double-glazing has additional benefits such as improved security and noise reduction that provide benefits in addition to reduced energy costs. Dwellings with a modern water heating system will have a more reliable system and/or a longer period to replacement in addition to reduced energy costs.

Third, usually, when reporting the price effect of an attribute, hedonic pricing studies present the result as a percentage price difference. In the case of EPC bands for instance, dwellings in band A may be reported as achieving prices that are 10% higher than those in band D. This is a relative measure and is usually expressed like this due to the way that the econometric models were constructed. However, if the benefit of energy efficiency is received largely in absolute terms, then this relative measure may distort the price effect. For example if homebuyers are prepared to pay an additional £100 for energy savings, in a location with typical sale prices of £1,000 per square metre, this would be 10% price premium but in a location with an average price of £5,000 per square metre, the same absolute price increase would represent a much smaller relative price premium of 2%. There is some empirical evidence to support this: a study for the Department of Energy and climate Change in England (discussed below) found notable regional variations in the size of relative price premiums..

3 Data and Modelling Approach

Building on the findings from the literature review we obtained dwelling-level data on market prices, building attributes and energy performance. Wales was chosen as the study area because the housing stock is relatively homogeneous in terms of age and building characteristics. Compared to England the proportion of flats and apartments is much lower and this allows the analysis to focus on detached, semi-detached and terraced houses. The data comprised a sample of 191,544 transactions that took place between 2 January 2003 and 26 February 2014. 47,158 (25%) of these included a second transaction and this sub-sample was used in a 'repeat sales' hedonic price model. The fields included in the sample, together with descriptive statistics, are listed in the appendix. Two attributes that are essential controls for any residential hedonic price modelling are size (represented by number of bedrooms in this data set) and age. The former was not recorded for 17% of the transactions and the latter 37%. Depending on the extent to which the same observations were missing both data items, this reduced the size of the sample for hedonic price modelling significantly. Whether there is any systematic bias in non-recording of number of bedrooms and age is not known.

Around a third of the transactions involved dwellings located in the Cardiff postcode area, and a quarter in the Swansea area. In all, around two thirds were in south Wales. For those transactions where the number of bedrooms was recorded (83% of the sample), 94% involved dwellings that had between two and four bedrooms. The vast majority of the sample observations involved freehold transactions. It would appear that, generally speaking, if a dwelling was built before the beginning of the twentieth century then it was recorded as having been built in 1900. This seems to be the logical explanation for nearly 16,000 dwellings being allocated to that year, a number comparable to the total built over the following 30 years. On that basis, and having taken the missing values into account, 19% of the sample transactions involved dwellings built before the twentieth century. Each decade from 1950 onwards comprises between 6% and 9% of the total number of transactions for which age has been recorded except 2000-09, which includes 15% of the sample. Each dwelling was geo-coded at the postcode level using the National Statistics Postcode Directory and this allowed further attributes to be appended including a demographic classification, an urban-rural indicator and a variable that recorded whether or not the dwelling was situated in a national park.

Regarding energy performance, in the UK an EPC assigns a rating to a dwelling on a scale of bands from A to G with A being the most efficient. The rating is based on energy relevant building characteristics including age, size, construction details, space and water heating, lighting and ventilation. Ratings are relative to standard energy use for the type of dwelling being assessed. New dwellings are assessed using the 2005 Standard Assessment Procedure (SAP) for Energy Rating of Dwellings. A reduced data SAP (RdSAP) is used to assess existing dwellings and this is of prime importance given the age of the vast majority of the housing stock. RdSAP consists of a series of defaults and inferences based on dwelling type, degree of detachment, age, and dimensions (see Table 1). Type and age of dwelling are used to infer window area. Wall type and age are used to infer U-values for walls, roofs and floors. Age is, therefore, a key variable.

Table 1 here

We employed a hedonic model to estimate the effect of EPC ratings on house prices. In the context of this research, Fuerst *et al.* (2015) have described the modelling approach in detail and in this paper we follow the approach closely. Housing as a differentiated commodity can be characterized by a vector *h* of various physical and locational attributes such as age (*a*), size (*f*), location (*L*), quality (*Q*) and EPC rating (*E*). A consumer draws utility variously from these components. Therefore, quantified measures of these attributes can provide a reasonable estimation of the consumer's utility function.

However, these individual attributes are not independent of each other and are not exogenously determined, which are what a regression model needs to assume to draw unbiased inferences. For example, we can easily imagine that EPC rating may be determined by a multitude of other confounding factors such as building material quality, design features, lighting, insulation, water heating and glazing. Moreover, current valuation of the benefits (S) to consumers may change and would depend on upon uncertain assumptions about future energy price inflation, behavioural patterns and appropriate discount rates. So, a property's price function can be empirically computed by the following equation:

$$P(h) = f(a, f, L, Q, E, S) \qquad (1)$$

Equation 1 can be put through standard regression techniques to estimate individual component effects. However, there are several caveats that need careful consideration before making inferences that are free of severe estimation biases.

First, the presence of correlation between Q and E is highly likely and, as a result, the coefficient estimate may either overstate or understate the true effect. Additionally, given the subjective and multi-faceted nature of the quality variable Q, there may be unobservable attributes that are correlated with the observed ones. Coupled with location dynamics, this presents a significant source of unobserved heterogeneity and omitted variable bias that can severely impair our ability to draw unbiased inferences. Second, typically energy efficient features tend to generate several direct and indirect and monetary and non-monetary benefits S. These benefits are difficult to quantify and, more importantly, variously contribute to the bias element in an estimation. Additionally, it becomes imperative to make some strong assumptions about future energy price inflation and appropriate discount rates, which also worsen the explanatory power of the estimation. Third, various types of consumer perceive energy efficiency features differently. More eco-friendly consumers may, for example, behave quite differently from other consumers. Such patterns in consumer behaviour remain un-accounted for in quantitative models. Finally, hedonic characteristics may impact on the price of a property in a non-linear fashion or in some cases the sign of the coefficient may reverse with an increase in the value of the hedonic regressor variable. For example, the price of a dwelling might respond to age in two ways; physical depreciation may reduce price paid but eventually price may respond positively to age – a 'vintage' effect. Moreover, older buildings tend to be less energy efficient and may involve significant refurbishment costs in order to comply with contemporary building regulations. Therefore, there may be a high correlation between age and EPC rating.

With above caveats in mind, we specify a hedonic price model as follows:

$$P_{it} = \propto_i + \sum_{i=1}^{I} \beta_i X_i + e_i \qquad (2)$$

Where P_{it} is the transaction price of a property (specified as the natural logarithm of price in £ per square metre), X_i is a vector of several explanatory locational and physical variables including our variable of interest, EPC rating; β_i is a vector of coefficients to be estimated; and e_i is a random error and stochastic disturbance term that is expected to take the form of a normal distribution with a mean of zero and variance of σ^2_e. We allow for temporal variation in the following form:

$$P_{it} = \sum_{j=1}^{J} \beta_j X_{jit} + \sum_{t=1}^{T} c_t D_t + e_{it} \qquad (3)$$

where c_t is the additional vector of estimated coefficients for each time period and D_t is a set of variables that takes the value of 1 if a house is sold in the period and 0 if it is not sold.

For the purpose of this study, we specify hedonic models to explain two dependent variables; price per square metre and price per square metre change (appreciation/depreciation). To capture the effects of EPC rating on these variables, we use a set of binary variables to indicate the EPC band of each dwelling at the relevant transaction date. Band D is the 'hold-out' category so the coefficients for the higher bands are expected to be positive. In addition to mitigating the effects of extreme values, as well as potential heteroskedasticity and non-normality, the semi-log specification of the hedonic model allows us to interpret the coefficients as average percentage premiums. However, we will also estimate a linear version of this hedonic model as a robustness check and to estimate the level of price premiums in pound sterling.

In the next step, we restrict the sample to dwellings for which records in our database indicate that an EPC was issued for the purpose of marketing the dwelling on the private rental market. While this identification of buy-to-let properties may not be perfect due to the possibility of switching between owner-occupation and leasing for some properties, we assume that the bulk of the properties thus identified are longer-term buy-to-let properties. To restrict the sample further, we only include those transactions in our sub-sample analysis which had an EPC issued before the sale of the dwelling, not earlier. This restriction is introduced to ensure that potential buy-to-let investors were aware of EPC performance when buying the property and were not forced to gather this information from other sources.

Finally, to measure the influence of EPC rating on price appreciation, we also undertake a hedonic analysis with the repeat sales transactions only. In doing so, we are able to exercise greater control over biases from the unit fixed effect as two sales of the same dwelling are compared, although, the assumption of no improvement or changes in property quality or other features is a concern. The repeat sales framework may take the following form:

$$P_i^2 - P_i^1 = \left(\sum_{j=1}^{J} \beta_j X_{ji}^2 + \sum_{t=1}^{T} c_t D_i^2\right) - \left(\sum_{j=1}^{J} \beta_j X_{ji}^1 + \sum_{t=1}^{T} c_t D_i^1\right) + e_i^{21} \qquad (4)$$

where the first and second sale periods are denoted by the superscripts 1 and 2 respectively. Equation (4) can be simplified to:

$$P_i^2 - P_i^1 = \sum_{t=1}^{T} c_t (D_i^2 - D_i^1) + e_i^{21} \qquad (5)$$

In our specification, we use a Wales house price index to capture 'expected' appreciation following the national trend as well as the property-specific price components in the following form:

$$\frac{P_t^2}{P_t^1} = \frac{I_t^2}{I_t^1} + \sum_{j=1}^{J} X_{jt} + u_j \qquad (6)$$

Thus price changes between two transactions are driven by the Wales-wide house price change, the time elapsed between the two sales and a set of observed and unobserved property characteristics that cause a house price to deviate from the national trend. We use the Land Registry (2014) house price index for Wales rather than calculating an index based on our own data sample as the former index is comprised of a much larger sample and hence reflects the market situation in Wales more accurately. The first factor is captured by the index ratio while the observed property-specific factors are represented by the vector of characteristics *X*. Finally, unobserved characteristics are captured in the error term *u*. Using this framework we are able to observe and estimate the magnitude of price differentials that result from dwellings being placed in different EPC bands.

4 Results

The descriptive statistics for the continuous variables in the sample are shown in the appendix. The dependent variable saleprice1psm (and saleprice2psm for repeat sales) was positively skewed so a log transformation was performed to normalize the distribution. Although the data set included 191,554 observations of dwelling transactions (and 47,158 repeat sales transactions), many did not include a complete set of attribute information; 23% did not have floor area recorded, 17% did not have the number of bedrooms recorded and 37% did not have any age information. This meant that the regression models operated on reduced sample sizes.

The independent variable of interest to this study is the EPC band. Over 85% of the dwellings in the sample are in Band C, D or E. The key control variables are floor area, age, property type, number of bedrooms and location. Of the total sample, 37% are terraced, 32% semi-detached, 29% detached and only 2% were flats. The dwelling stock is quite old; 71% of the dwellings were built before 1960 (with 40% built before 1900) and fewer than 10% were constructed between 2000 and 2009. 69% of the sample was located in an urban setting. Figure 1 illustrates the relationship between property type, EPC band and mean sale price. The price differential between detached, semi-detached and terraced houses and flats is as expected; detached dwellings selling for the highest prices and flats the lowest. It is also possible to discern a relationship between price and EPC band, particularly in the case of flats but also, to a less obvious extent, in the case of semi-detached and terraced dwellings. Notwithstanding the point made above in relation to the very small number of observations in bands A, B, F and G, there does not appear to be a discernible relationship between EPC band and price paid for detached dwellings.

Figure 1 here

Figure 2 shows the geographical distribution of EPC ratings for the sample of dwellings. The map was created by interpolating a surface (using an inverse distance weighting algorithm) from the location and EPC ratings of the sample of dwellings. The lighter shading represents areas with relatively high EPC ratings and the darker areas are where dwellings have generally lower EPC ratings. The urban areas, particularly in south Wales, contain greater concentrations of dwellings with higher EPC ratings and dwellings with generally lower ratings dominate the more rural locations in the north west of Wales.

Figure 2 here

After removing observations with missing variables and selecting transactions that had taken place from the beginning of 2008 onwards (84,776 observations), the cross-section regression sample comprised 62,464 observations and the repeat sales regression sample included 25,189 observations.

When analysing the price effect by property type these sample sizes reduce accordingly and are reported at the bottom of the results tables.

The determinants of price per square metre

Following the modelling approach and data sampling outlined in section three, we first fit regression models to both the full set of observations and the sub-samples of the different types of dwelling. The results are presented in Table 2. Price per square metre is easier to interpret than the more commonly used total price as it eliminates the size effect from the dependent variable. This size effect may generate undesirable effects as it tends to inflate the predictive power of a hedonic model (for example, our baseline R^2 would increase from 0.50 to 0.74 when total price is used) and may hence swamp the effects of less powerful predictors of price in the hedonic model. The (log of) house price per square metre is explained as a function of four dwelling attributes (age, dwelling type, number of bedrooms and tenure), a neighbourhood attribute (urban-rural index score) and EPC rating/point. The fact that housing transactions took place in different time periods and different areas is controlled for by including quarterly time fixed effects and postcode area fixed effects in the model.

The overall explanatory power of the model is good with an adj. R^2 of around 51% for the full sample. The coefficients of the explanatory variables largely have the expected signs. With regard to the EPC rating, using band D as the 'hold-out' category, the pattern of price effects is consistent with a positive relationship between energy performance rating and sale price. Adjusting the coefficients of the semilogarithmic estimation for potential bias as proposed by Kennedy (1981) and van Garderen & Shah (2002) yields minimally different effects in the order of magnitude of 0.01%. For the whole sample estimation, there are significant positive premiums for dwellings in bands A/B (11.3%) and C (2.1%) compared to dwellings in band D. For dwellings in EPC bands lower than D there are statistically significant discounts; -2.1% for band E dwellings, -4.7% for band F dwellings and -7.2% for dwellings in band G. The price impact varies depending on the type of property: a terraced dwelling rated B sold for approximately 17.1% more per square metre than a terraced dwelling EPC rated D. The comparable figure for a semi-detached dwelling is 8.2%. Relative to the other dwelling types, detached dwellings are likely to display the greatest degree of heterogeneity, particularly in rural areas. Recognising this, detached dwellings were categorised as urban or rural. Table 2 shows that the price impact is more marked and for urban dwellings in bands E and F than for rural dwellings in the same bands. This might be a result of purchasers' willingness to pay higher prices for rural dwellings (perhaps because of their character and setting) regardless of their energy performance. In the last column of Table 2 the results of the estimation when energy efficiency score, rather than band, is used as the independent variable are displayed. The expected positive relationship between energy efficiency and dwelling sale price is also found.

Table 2 here

Turning to the control variables, all else equal, each additional bedroom will lower the unit price of a property, possibly reflecting decreasing marginal utility of each square metre of living space consumed. An additional effect working in the same direction might be that dwellings of a given size will achieve lower prices if they are divided up into more rooms, i.e. a discount on properties with small rooms. The effect of age on dwelling price per square metre is non-linear and variable between dwelling types. Compared to dwellings constructed pre-1900, dwellings constructed since 1983 have sold for small but statistically significant price premiums. When we look at the results across dwelling types, it is apparent that there are notable differences between semi-detached, terraced and detached properties. In contrast to semi-detached and terraced dwellings, detached dwellings constructed between 1900 and 2003 tend to sell for significantly less per square metre than dwellings constructed before 1900. Regarding dwelling type the results are in line with expectations; with terraced dwellings as the 'hold-out' category, flats, semi-detached and detached properties achieve significantly higher prices per square metre, with the latter selling for approximately 28% more per square metre than terraced dwellings. The coefficients for the urban-rural indexes also have the expected signs. Compared to leasehold, the coefficient for freehold is insignificant.

The private rental market

We now turn to the private rented sector of the housing market in Wales. As mentioned in the literature review, a difference in the quality of the stock between owner-occupied and private rented dwellings may introduce bias in our estimation results, particularly if some of the quality characteristics are unobserved and are correlated with the EPC ratings. However, the summary statistics suggests that this is not a major concern for the present study of Wales. The average sale price in our private rental subsample is very similar to the overall average price (£142,000 and £145,000 respectively). In terms of energy efficiency ratings, we find a similar distribution of EPC bands and scores. The average EPC score in the private rental segment is 57.3 compared to 58.3 in the overall sample. To further mitigate the potential for any omitted variables bias due to unobserved quality differences, we conduct the estimation of the subsample separately rather than including interaction terms in the main model. Hence, the reference point for EPC band capitalisation is an average D-rated rental property, rather than a standard D-rated property of either tenure status.

Table 3 reveals that energy efficient dwellings in bands A, B and C achieve price premiums that are comparable to the general market. This is to be expected in a market setting where buy-to-let landlords compete with owner-occupiers for these properties. However, we do not find significant discounts for rental dwellings with below average energy efficiency ratings. The implications of these findings are discussed in more detail in the following section.

Table 3 here

The determinants of price appreciation per square metre

We apply a similar regression specification with dwelling price *appreciation* per square metre as the dependent variable. We do not have definite prior expectations for either positive or negative effects.

It is possible that price premiums associated with superior energy performance have been factored into initial prices and that there is no ‘growth premium’. On the other hand, it is possible that the increasing awareness of energy and environmental issues in the last decade has meant that price effects have produced positive effects on price appreciation. In other words, the effects of superior energy performance on initial prices may be positive and, due to subsequent greater demand for energy efficient dwellings, the effects on price appreciation may also be positive.

Table 4 provides estimates of the determinants of dwelling price appreciation per square metre. The results are not as consistent as the price model. C-rated dwellings have experienced significantly higher price appreciation than D-rated dwellings. However, this is not the case for the dwellings in the A/B bands, which have experienced no statistically significant higher price appreciation than D-rated dwellings. Dwellings rated E and F are estimated to have grown at statistically significant lower rates compared to D-rated properties. When we look at the estimates for dwelling type sub-samples, the effects on price appreciation are largest for terraced dwellings with no significant effect on flats. There is a significant and positive price impact for detached dwellings in urban areas in band C, and for dwellings in bands E and F the negative price growth impact is also significant.

Table 4 here

<u>Robustness checks</u>

As noted above, a common issue with hedonic estimations is potential omitted variable bias. In the context of this research one particular concern is that houses with better EPC ratings may have been subject to unobserved improvements that enhance the quality of dwellings in addition to enhancing energy performance. In order to try to counteract such potential bias, we run the models with a number of restricted samples. The main purpose of the restrictions is to exclude dwellings that are more likely to have been improved or that may be unusual in some way e.g. dwellings that have been re-sold within a short period of time or dwellings exhibiting high levels of unexplained variance in the estimations. In order to try to eliminate the effects of potential unobserved changes to houses to bias the estimates, we restrict the sample to houses built relatively recently (since 2000).

The results are presented in Table 5 and it is reassuring to see that they remain broadly stable. For the cross-sectional price models, the results of the restricted sample models estimate similar patterns of premiums and discounts compared to EPC band D as the full sample model. Whilst restricting the sample to dwellings with low unexplained variance reduces the sample to 9,866 sale transactions, the estimates of price premiums and discounts remain indistinguishable from the full sample results. When the sample is restricted to the 9,851 dwellings that have been built and sold since 2000, the results do change. Compared to a band D dwelling, the estimated price premium for band A/B dwellings drops to 4.5%. The price premium for band C dwellings decreases to less than 1%. Bearing in mind that only a small proportion of modern houses will have energy ratings below D, we find no significant discount for poor energy efficiency performance. Similarly, excluding dwellings that have been sold twice or ‘flipped’ in under two years has no notable effect on the estimated coefficients. Applying similar restrictions to the price appreciation model results in a similar pattern with little variation in the estimated coefficients for the various restricted samples. The exception is EPC band G where the similar-sized effect has taken on a degree of statistical significance.

Table 5 here

A further robustness check concerns the functional form of the hedonic equation. The main semi-log model used in our analysis has a number of advantages as outlined in the methods section above but the estimated coefficients are expressed as a percentage of the overall transaction price, so a fixed amount will be lower in high house price areas compared to lower priced areas. To neutralise the effect of the variation in the underlying denominator (the house price), we estimate a linear version of the hedonic model where the dependent variable is expressed in levels rather than in logarithmic form. The results reported in Table 6 show that the estimated coefficients are largely in line with the estimated percentage values of the semi-log form. Additionally, the linear estimation shows that the highest EPC bands A and B add £230 to the square metre price compared to the average D category and Band C adds approximately £40 to the average property. Conversely, the discounts to properties achieving below average energy ratings are confirmed. The breakdown by property type underlines that these premiums and discounts are strongest in the semi-detached and terraced submarkets and less pronounced for detached houses and flats. However, the semilog hedonic model remains our main specification in this study as the percentage estimates obtained from it facilitate comparison across studies that use different monetary units and currencies.

Table 6 here

5 Discussion and implications of the findings

There are several reasons for expecting house buyers to pay more for an energy efficient house relative to a very similar house that is less energy efficient. Lower energy bills essentially result in higher household disposable income. The Building Research Establishment found, for a sample of 125 dwellings that represented a broad range of house types, a strong correlation between annual energy costs and EPC rating/band (BRE, 2014). The presence of high quality water heating equipment, lighting, etc. should reduce expenditure on replacement and maintenance. Some house buyers may obtain a psychic income from eco-consumption. There can be additional benefits from energy efficient features that have little to do with energy efficiency. For instance, double-glazing reduces noise pollution and increases security. In short, there are grounds to expect a positive price effect of energy efficiency, all else equal. However, the relative effect is not expected to be uniform.

In order to isolate and estimate the effect of the EPC, it is important to be able to take into account the other factors that are affecting the price of the house - the date of sale, its condition, location, age, size, type, quality of fittings, etc. Such comprehensive data is rarely available and researchers trying to isolate and identify house price determinants tend to be concerned about omitted variable problems. A particular concern in this study is that an unobserved variable such as condition, quality or recent improvements may be related to energy efficiency and consequently their (unobserved) effect on house price may be misattributed as being due to energy efficiency. Missing variables that are not linked to energy efficiency (aspect, view and proximity to busy road for example) may affect house prices and reduce the explanatory power of the statistical model but they will not bias the estimations. In this study we have tried to reduce the risk of this type of problem by removing houses from the sample that are more likely to have been improved or have better quality fittings.

Regarding the determinants of price per square metre, and the influence of energy performance in particular, the estimated price premiums are much higher than for the comparable study conducted in England (Fuerst *et al*, 2015). One reason for this is the lower average house price in Wales. The findings for Wales are very similar to the results for the North East region of England where significant positive premiums were estimated for dwellings in bands A and B (14.4%) or C (2.7%) compared to dwellings in band D and statistically significant discounts for dwellings in band E (-2.5%) and F (-6.0%).

The findings for the effect of EPC rating on house price growth are less consistent. Compared to band D dwellings, those in band C have experienced significantly higher house price growth. Surprisingly this is not found for dwellings in bands A or B, which have experienced no statistically significant higher price appreciation than D-rated dwellings. The DECC study in England actually found significantly lower rates of growth for this category. There is evidence of a negative effect of poor energy efficiency on house price growth. Dwellings in bands E and F are estimated to have grown at statistically significant lower rates compared to band D dwellings.

We investigated a subsample of properties which had an EPC issued for the purpose of a private rental marketing. The absence of significant discounts for rental dwellings with below average energy ratings may be taken as further evidence of the split incentive problem. Buy-to-let investors may not apply the same discounts to low-EPC dwellings that owner-occupiers would because energy costs are passed on to tenants. Given that rental premiums paid by tenants are uncertain, buy-to-let buyers may outbid owner-occupiers in this segment of the market. As a consequence, we do not observe a significant discount for lower-rated properties. This finding is in line with the previous literature, in particular Hyland *et al.* (2013) who find that the rental premium captures only 14-55% of the net present value of energy savings. Rehdanz (2007) and Kholodilin and Michelsen (2014) arrive at similar conclusions in their studies of German housing markets. The implicit lower return on energy efficiency for landlords compared to owner occupiers thus leads to a levelling of prices between D, E, F and G bands, all else equal. A diverging result compared to the German studies is our finding of a significant premium for A, B, C rated properties which may be explained by the fact that the owner-occupied and rental tiers of the market are less segmented in the UK market and the Welsh market in particular. The fraction of 'dedicated' rental stock on the overall market is lower and most properties could be used for either owner occupation or as a rental investment, which is not necessarily the case in Germany.

6 Conclusion

The introduction of mandatory energy performance ratings for the commercial and residential real estate sectors across the European Union and in many other countries reflects a growing focus on reducing carbon emissions from the real estate stock. The main objective of energy efficiency certification is to influence the behaviour of consumers. In the context of the residential housing market EPCs are intended to provide trustworthy information to house buyers about energy efficiency. It is expected that energy savings associated with energy efficient attributes such as insulation, double/triple glazing, efficient water heating, low energy lighting, etc. will lead to house price premiums which will, in turn, lead to increased adoption of energy efficient features. However, it is important to acknowledge that untangling and isolating the effect of a single variable on the price of a house presents methodological challenges.

Drawing upon a large sample of house sales, we find a positive association between dwelling price per square metre and energy performance rating. Almost certainly due to a lower average house price, these estimated price premiums are much higher than for the comparable study in England. It is notable that the findings for Wales are very similar to the results for the North East region of England. There are statistically significant positive price premiums for dwellings in bands A/B (12.8%) and C (3.5%) compared to houses in band D. For dwellings in band E (-3.6%) and F (-6.5%) there are statistically significant discounts. The relative price effects are highest for terraced dwellings. In order to try to take account of dwellings that may have been improved or be of better quality or condition, we also excluded dwellings whose prices were being poorly explained by the econometric model. The pattern of price effects remains broadly unchanged. In line with expectations, EPC C-rated dwellings experienced a higher rate of house price growth than D-rated dwellings whereas E and F-rated dwellings experienced lower rates.

Finally, our finding of no discounts for dwellings in bands E, F and G adds to the emerging evidence of the split incentive problem and its impact on transaction prices in the private rental segment of the market. Incoming legislation in the UK which forbids the leasing of dwellings below a minimum energy efficiency rating from 2017 onwards may alter the price patterns for low-energy efficiency properties, which warrants a follow-up study that could also model the relationship between prices, rents and energy bills more explicitly than this paper was able to do with information on sales transaction prices only.

References

Australian Bureau of Statistics (2008) *Energy Efficiency Rating and House Prices in the ACT*, Report for Department of the Environment, Water, Heritage and Arts.

BRE (2014) EPCs and mortgages: demonstrating the link between fuel affordability and mortgage lending, Wales Low Zero Carbon Hub / Constructing Excellence in Wales, http://www.cewales.org.uk/cew/wp-content/uploads/EPCs-Mortgages1.pdf.

Brounen, D. and Kok, N. (2011) On the Economics of Energy Labelling in the Housing Market, *Journal of Environmental Economics and Management*, 62, 166-179.

Cajias, M., & Piazolo, D. (2013) Green performs better: energy efficiency and financial return on buildings, *Journal of Corporate Real Estate*, *15*(1), 53-72.

Deng, Y. and Wu, J. (2014) Economic returns to residential green building investment: the developers' perspective, *Regional Science and Urban Economics*, 47, 35-44.

Dinan, T. and Miranowski, J. 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.

European Commission (2013) *Energy performance certificates in buildings and their impact on transaction prices and rents in selected EU countries*, Final Report, European Commission (DG Energy), 19 April.

Feige, A., McAllister, P. and Wallbaum, H. (2013) Rental price and sustainability ratings: which sustainability criteria are really paying back? *Construction Management and Economics*, 31(4), 322-334.

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, 46-69.

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, 145-156.

Fuerst, F., van de Wetering, J. and Wyatt, P. (2013) Is intrinsic energy efficiency reflected in the pricing of office leases? *Building Research & Information*, 41(4), 373-383.

van Garderen, K. J. and Shah, C. (2002) Exact interpretation of dummy variables in semilogarithmic equations, *Econometrics Journal*, 5, 149-159

Halvorsen, R. and Pollakowski, H. 1981 The effects of fuel prices on house prices, *Urban Studies*., 18, 205–211.

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

Iwata, S., Yamaga, H., (2008) Rental externality, tenure security, and housing quality. *Journal of Housing Economics* 17 (3), 201-211.

Johnson, R. and Kaserman, D. 1983 Housing Market Capitalization of Energy-Saving Durable Good Investments. *Economic Inquiry*, 21(3), 374-386.

Kahn, M. and Kok, N. (2014) The capitalisation of green labels in the California housing market, *Regional Science and Urban Economics*, 47, 25-34.

Kennedy, P. E. (1981) Estimation with correctly interpreted dummy variables in semilogarithmic equations, *American Economic Review*, 71(4).

Kholodilin, K. A., Michelsen, C. (2014): The Market Value of Energy Efficiency in Buildings and the Mode of Tenure, Discussion Papers of DIW Berlin 1398, DIW Berlin, German Institute for Economic Research.

Kok, N. and Jennen, M. (2012) The impact of energy labels and accessibility on office rents, *Energy Policy*, 46, 489-497.

Land Registry (2014): *House Price Index, 1995-2014*. Accessed September 2014.

Laquatra, J. 1986. Housing market capitalization of thermal integrity. *Energy Economics*, 8(3), 134-138.

Laquatra, J., Dacquisto, D. Emrath, P. and Laitner, J. 2002. Housing Market Capitalization of Energy Efficiency Revisited. *ACEEE Summer Study of Energy Efficiency in Buildings, Teaming for Efficiency, Proceedings, 8*.

Quigley, J. 1984. The production of housing services and the derived demand for residential energy. *The RAND Journal of Economics*, 555-567.

Quigley, J. and Rubinfeld, D. 1989. Unobservables in consumer choice: Residential energy and the demand for comfort. *The Review of Economics and Statistics*, 416-425.

Rehdanz, K. (2007) Determinants of residential space heating expenditures in Germany. *Energy Economics* 29 (2), 167-182.

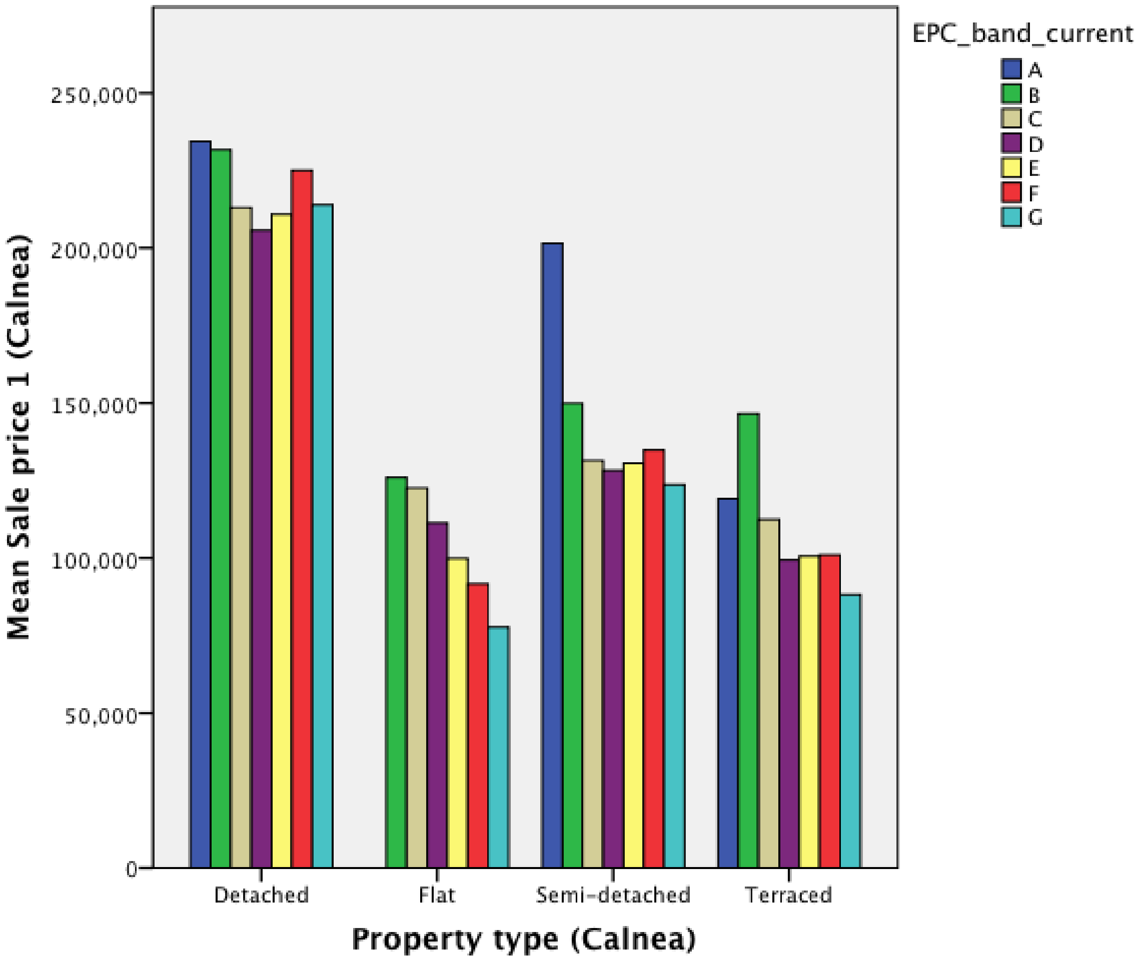

Figure 1: Relationship between property type, EPC band and sale price

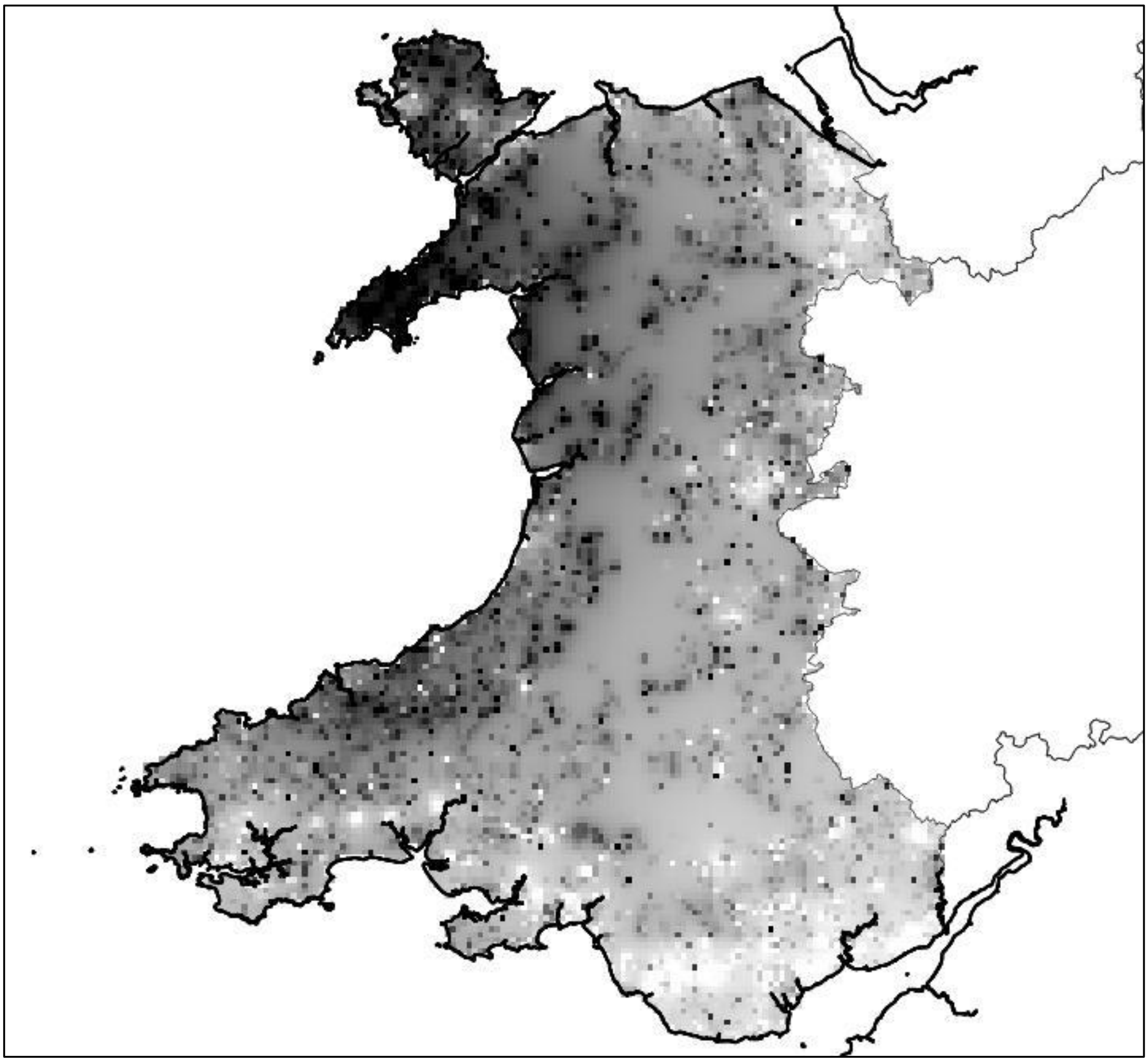

Figure 2: EPC ratings (darker areas low rating, lighter areas high rating)

Table 1: RdSAP dwelling categorisations

Type	Degree of detachment	Age band
House	Detached	Before 1900
Bungalow	Semi-detached	1900-1929
Flat	Mid-terrace	1930-1949
Maisonette	End-terrace	1950-1964
	Enclosed mid-terrace	1965-1975
	Enclosed end-terrace	1976-1982
		1983-1990
		1991-1995
		1996-2002
		2003-2006
		2006-

Table 2: Energy rating and price: hedonic estimations
(dependent variable: log of price per square metre)

	Full sample (EPC bands)	Detached	Detached (rural)	Detached (urban)	Semi-detached	Terraced	Flat	Full sample (EPC rating)
EPC band A/B	0.113*** (11.37)	-0.0199 (-1.20)	-0.0181 (-0.52)	-0.0200 (-1.10)	0.0824*** (4.76)	0.171*** (8.80)	0.0355 (0.76)	
EPC band C	0.0206*** (6.10)	0.00197 (0.33)	-0.00155 (-0.14)	0.00274 (0.40)	0.00395 (0.73)	0.0234*** (3.76)	0.0388 (1.46)	
EPC band D	Hold-out	Hold-out	Hold-out	Hold-out	Hold-out	Hold-out	Hold-out	
EPC band E	-0.0209*** (-6.76)	-0.0174** (-2.72)	-0.00580 (-0.55)	-0.0214** (-2.67)	-0.0204*** (-3.97)	-0.0361*** (-7.63)	-0.0824 (-1.86)	
EPC band F	-0.0473*** (-8.58)	-0.0442*** (-4.45)	-0.0305* (-2.16)	-0.0687*** (-4.90)	-0.0551*** (-5.60)	-0.0945*** (-10.70)	-0.105 (-1.30)	
EPC band G	-0.0717*** (-6.90)	-0.0499** (-2.78)	-0.0591** (-2.76)	-0.0527 (-1.44)	-0.0832*** (-4.11)	-0.140*** (-8.55)	-0.150 (-1.83)	
EPC rating								0.0432*** (11.21)
Number of bedrooms	-0.0601*** (-33.93)	-0.0557*** (-19.34)	-0.0721*** (-15.67)	-0.0431*** (-12.20)	-0.0382*** (-11.36)	-0.0529*** (-16.90)	-0.0489* (-2.27)	-0.0607*** (-34.27)
1900	-0.142*** (-21.14)	-0.174*** (-9.24)	-0.133*** (-5.58)	-0.130*** (-3.86)	-0.150*** (-9.22)	-0.0811*** (-9.87)	-0.0903 (-1.14)	-0.145*** (-21.47)
1901-29	-0.106*** (-15.62)	-0.138*** (-8.00)	-0.113*** (-4.74)	-0.0442 (-1.43)	-0.0851*** (-5.56)	-0.0565*** (-6.67)	-0.0855 (-1.32)	-0.108*** (-15.90)

1930-49	-0.0792***	-0.0812***	-0.108***	0.0469	-0.0626***	-0.0631***	0.0940	-0.0796***
	(-11.14)	(-5.26)	(-4.32)	(1.63)	(-4.31)	(-5.64)	(1.07)	(-11.17)
1950-59	-0.141***	-0.0457**	-0.0879***	0.0842**	-0.145***	-0.143***	-0.297*	-0.140***
	(-19.41)	(-2.87)	(-3.66)	(2.86)	(-9.88)	(-12.94)	(-2.27)	(-19.36)
1960-69	-0.0594***	-0.0550***	-0.0805***	0.0659*	-0.0172	-0.150***	-0.141*	-0.0586***
	(-8.12)	(-3.87)	(-4.38)	(2.32)	(-1.14)	(-12.96)	(-2.07)	(-7.99)
1970-79	-0.0127	-0.0814***	-0.107***	0.0385	0.0657***	-0.0796***	-0.0891	-0.0105
	(-1.78)	(-6.08)	(-6.59)	(1.37)	(4.37)	(-6.69)	(-1.27)	(-1.48)
1980-89	0.0509***	-0.0603***	-0.0853***	0.0589*	0.112***	0.128***	-0.00250	0.0552***
	(6.86)	(-4.37)	(-5.03)	(2.07)	(7.14)	(10.53)	(-0.04)	(7.45)
1990-99	0.0893***	-0.0373**	-0.0667***	0.0796**	0.148***	0.219***	-0.0111	0.0959***
	(12.65)	(-2.77)	(-3.95)	(2.85)	(9.72)	(20.01)	(-0.19)	(13.71)
2000-09	0.0867***	-0.0177	-0.0551**	0.105***	0.162***	0.191***	0.0530	0.109***
	(11.53)	(-1.27)	(-3.07)	(3.69)	(10.04)	(14.80)	(0.93)	(14.95)
2010-	0.110***	0.0205	-0.0341	0.155***	0.176***	0.219***	0.208**	0.172***
	(10.35)	(1.14)	(-1.22)	(4.96)	(8.58)	(10.24)	(3.00)	(19.42)
Detached	0.277***							0.272***
	(69.75)							(68.98)
Semi-detached	0.127***							0.126***
	(39.08)							(38.68)
Terraced	Hold-out							Hold-out
Flat	0.0406**							0.0449**

	(2.77)							(3.05)
Tenure	0.0174*	0.0235	0.0425	0.0119	0.0133	0.0272*	-0.392**	0.0166*
	(2.33)	(1.59)	(1.40)	(0.75)	(1.14)	(2.13)	(-3.07)	(2.23)
Urban-rural indicator	0.00111	0.0316***			-0.0138*	-0.0244***	0.206*	0.00000389
	(0.28)	(4.78)			(-2.03)	(-3.30)	(2.31)	(0.00)
Constant	7.422***	7.712***	7.770***	7.584***	7.446***	7.376***	7.799***	7.247***
	(512.80)	(236.90)	(123.94)	(183.50)	(285.27)	(339.82)	(85.07)	(352.01)
Quarterly fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Postcode fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
adj. R-sq	0.505	0.260	0.201	0.343	0.429	0.571	0.518	0.504
N	62,464	18,568	7,686	10,882	21,069	22,109	718	62,461

t statistics in parentheses * p<0.05, ** p<0.01, *** p<0.001

Table 3: Energy rating of private rental properties and price: hedonic estimations (dependent variable: log of price per square metre)

	Private rentals
EPC band_A/B	0.185**
	(3.11)
EPC band_C	0.040*
	(2.27)
EPC band_E	-0.022
	(-1.55)
EPC band_F	-0.017
	(-0.60)
EPC band_G	-0.072
	(-1.41)
Full set of controls	Y
adj. R^2	0.497
N	3,182

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Energy rating and price: hedonic estimations
(dependent variable: house price appreciation per square metre)

	Full sample (EPC bands)	Detached	Detached (rural)	Detached (urban)	Semi-detached	Terraced	Flat
EPC band A/B	-0.00169	-0.00580	-0.0195	0.0266	-0.0351	0.0798^{**}	-0.0154
	(-0.11)	(-0.15)	(-0.21)	(0.71)	(-1.71)	(2.96)	(-0.33)
EPC band C	0.0322^{***}	0.0326^{***}	0.0493^{*}	0.0319^{**}	0.0114	0.0505^{***}	-0.0159
	(5.44)	(3.37)	(2.45)	(3.06)	(1.21)	(4.12)	(-0.50)
EPC band D	Hold-out	Hold-out	Hold-out	Hold-out	Hold-out	Hold-out	Hold-out
EPC band E	-0.0449^{***}	-0.0489^{***}	-0.0114	-0.0776^{***}	-0.0336^{**}	-0.0518^{***}	0.0248
	(-7.40)	(-3.62)	(-0.47)	(-5.01)	(-3.28)	(-5.60)	(0.43)
EPC band F	-0.0591^{***}	-0.122^{***}	-0.0745^{**}	-0.185^{***}	-0.0720^{***}	-0.0205	0.101
	(-5.35)	(-6.05)	(-2.70)	(-6.53)	(-3.40)	(-1.12)	(0.90)
EPC band G	-0.0153	-0.104^{*}	-0.0539	-0.208^{**}	-0.0174	0.0369	-0.251
	(-0.62)	(-2.35)	(-1.01)	(-2.63)	(-0.35)	(0.89)	(-1.47)
House price index	1.215^{***}	0.977^{***}	1.021^{***}	0.938^{***}	1.277^{***}	1.351^{***}	0.695^{***}
	(70.89)	(35.13)	(19.47)	(29.54)	(42.38)	(44.36)	(6.03)
No. of beds	0.00913^{**}	0.0141^{**}	0.0188^{*}	0.0110	0.00874	0.0110^{*}	-0.0490
	(3.10)	(2.60)	(1.99)	(1.66)	(1.73)	(2.06)	(-1.79)
1900	0.0305^{**}	0.0211	0.00723	0.0375	0.0115	0.0351^{*}	-0.298^{*}
	(2.67)	(0.60)	(0.16)	(0.56)	(0.41)	(2.53)	(-2.08)
1901-29	0.0145	-0.0229	-0.0394	-0.0400	0.0286	0.0122	-0.195
	(1.28)	(-0.76)	(-0.95)	(-0.71)	(1.07)	(0.86)	(-1.33)
1930-49	0.0119	-0.0161	-0.0412	-0.0399	0.0217	-0.00511	-0.338
	(0.96)	(-0.58)	(-0.84)	(-0.79)	(0.83)	(-0.26)	(-1.86)

1950-59	0.0427**	0.0459	0.119	-0.0226	0.0438	0.0436	-0.259
	(2.99)	(1.21)	(1.50)	(-0.40)	(1.61)	(1.73)	(-1.91)
1960-69	-0.00770	-0.0225	-0.00407	-0.0650	-0.0133	0.000116	-0.219
	(-0.60)	(-0.78)	(-0.09)	(-1.26)	(-0.50)	(0.01)	(-1.64)
1970-79	-0.0580***	-0.113***	-0.110***	-0.141**	-0.0430	-0.0385	-0.411**
	(-4.91)	(-4.45)	(-3.33)	(-2.78)	(-1.69)	(-1.84)	(-2.95)
1980-89	-0.0887***	-0.135***	-0.111***	-0.184***	-0.0640*	-0.102***	-0.310*
	(-7.50)	(-5.21)	(-3.41)	(-3.55)	(-2.39)	(-5.86)	(-2.49)
1990-99	-0.118***	-0.167***	-0.153***	-0.214***	-0.101***	-0.0972***	-0.375**
	(-10.33)	(-6.69)	(-4.73)	(-4.28)	(-3.88)	(-5.73)	(-2.93)
2000-09	-0.145***	-0.179***	-0.137***	-0.239***	-0.124***	-0.158***	-0.437***
	(-12.01)	(-6.86)	(-3.97)	(-4.67)	(-4.69)	(-8.69)	(-3.40)
2010-	-0.139***	-0.131*	0.0676	-0.276***	-0.127***	-0.236***	
	(-4.77)	(-2.12)	(0.48)	(-4.38)	(-3.37)	(-8.05)	
Detached	0.0921***						
	(4.79)						
Semi-detached	0.0844***						
	(4.52)						
Terraced	0.0925***						
	(5.00)						
Flat	Hold-out						
Freehold	0.0109	0.0181	0.0420	0.00857	0.0257	-0.00260	-0.520***
	(0.95)	(1.05)	(1.36)	(0.39)	(1.49)	(-0.12)	(-5.45)
Urban-rural indicator	0.0100	0.0285*			-0.00992	0.0218	0.0643
	(1.31)	(2.16)			(-0.79)	(1.42)	(0.67)
Constant	0.0895***	0.185***	0.144**	0.254***	0.155***	0.190***	0.434**

	(4.78)	(5.25)	(2.74)	(4.44)	(4.79)	(6.71)	(3.12)
Postcode fixed effects	Y	Y			Y	Y	Y
adj. R^2	0.256	0.213	0.179	0.244	0.287	0.263	0.422
N	25,189	6,971	2,600	4,371	8,066	9,813	339

t statistics in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Robustness checks - model results with restricted samples

	House price models		House price appreciation models
Sample Restriction	Built since 2000	Residual within 0.05	Winsorized residual
EPC band A/B	0.0454***	0.112***	-0.00441
	(3.82)	(53.12)	(-1.51)
EPC band C	0.00757	0.0196***	0.0294***
	(1.04)	(22.38)	(26.78)
EPC band D	Hold-out	Hold-out	Hold-out
EPC band E	-0.0914**	-0.0215***	-0.0438***
	(-3.13)	(-27.23)	(-36.07)
EPC band F	-0.110**	-0.0478***	-0.0598***
	(-2.82)	(-36.72)	(-26.02)
EPC band G	-0.0868	-0.0724***	-0.0150**
	(0.58)	(-31.71)	(-2.99)
Adj. R^2	0.230	0.990	0.975
N	9851	9866	5414

t statistics in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Energy rating and price: hedonic estimations
(dependent variable: price per square metre)

	Full sample (EPC bands)	Detached	Detached (rural)	Detached (urban)	Semi-detached	Terraced	Flat	Full sample (EPC rating)
EPC band A/B	230.8*** (10.58)	8.755 (0.25)	3.057 (0.05)	13.01 (0.32)	207.2*** (5.44)	302.5*** (6.56)	126.2 (1.30)	
EPC band C	39.78*** (6.46)	25.47 (1.78)	2.828 (0.11)	36.68* (2.09)	10.18 (1.17)	32.82*** (3.51)	86.49 (1.51)	
EPC band D	Hold-out	Hold-out	Hold-out	Hold-out	Hold-out	Hold-out	Hold-out	
EPC band E	-16.80*** (-3.32)	-25.79 (-1.92)	-19.55 (-0.85)	-25.20 (-1.57)	-20.90* (-2.56)	-33.48*** (-5.14)	-119.9 (-1.56)	
EPC band F	-33.23*** (-3.59)	-61.40** (-2.99)	-55.40 (-1.78)	-86.05*** (-3.31)	-40.74** (-2.63)	-91.45*** (-8.08)	-147.5 (-1.02)	
EPC band G	-37.41 (-1.94)	-44.15 (-1.07)	-70.20 (-1.36)	-54.53 (-0.77)	-70.45* (-2.22)	-122.4*** (-5.42)	-357.7* (-2.22)	
EPC rating								38.92*** (5.51)
Number of bedrooms	-91.79*** (-28.57)	-96.85*** (-15.77)	-127.4*** (-12.24)	-73.28*** (-10.40)	-56.51*** (-9.69)	-70.66*** (-14.32)	-79.95 (-1.79)	-92.64*** (-28.88)
1900	-220.3*** (-18.72)	-324.6*** (-8.74)	-265.7*** (-5.42)	-218.4*** (-3.59)	-244.4*** (-9.21)	-112.6*** (-9.05)	-88.62 (-0.56)	-226.2*** (-19.07)
1901-29	-184.0***	-273.6***	-251.2***	-78.93	-157.6***	-90.11***	-137.3	-189.5***

	(-15.27)	(-7.65)	(-5.14)	(-1.29)	(-5.99)	(-7.10)	(-1.07)	(-15.62)
1930-49	-166.0***	-213.9***	-259.9***	29.70	-144.3***	-114.1***	163.3	-169.4***
	(-12.77)	(-6.50)	(-4.83)	(0.52)	(-5.66)	(-6.90)	(0.88)	(-12.93)
1950-59	-253.4***	-146.5***	-249.9***	114.9	-266.3***	-226.9***	-504.4*	-256.0***
	(-19.30)	(-4.25)	(-5.04)	(1.93)	(-10.40)	(-13.94)	(-2.47)	(-19.34)
1960-69	-136.2***	-186.7***	-240.0***	51.66	-77.51**	-223.1***	-285.9*	-137.8***
	(-10.05)	(-5.95)	(-5.68)	(0.91)	(-2.94)	(-13.57)	(-2.09)	(-10.11)
1970-79	-84.59***	-249.4***	-302.2***	-11.11	34.48	-147.0***	-168.9	-84.31***
	(-6.29)	(-8.40)	(-8.29)	(-0.20)	(1.29)	(-8.09)	(-1.09)	(-6.24)
1980-89	21.26	-212.5***	-265.6***	23.96	113.1***	191.0***	9.382	24.69
	(1.47)	(-6.85)	(-7.12)	(0.41)	(4.01)	(9.43)	(0.07)	(1.71)
1990-99	65.94***	-193.5***	-243.6***	34.51	158.5***	321.0***	-14.52	73.39***
	(4.88)	(-6.48)	(-6.52)	(0.61)	(5.86)	(18.10)	(-0.12)	(5.48)
2000-09	66.72***	-156.7***	-218.1***	78.72	199.4***	278.0***	85.25	108.4***
	(4.55)	(-4.99)	(-5.27)	(1.37)	(6.83)	(12.10)	(0.70)	(7.65)
2010-	82.16***	-81.22	-180.5**	178.0**	158.5***	273.7***	417.6**	212.1***
	(3.56)	(-1.91)	(-2.73)	(2.71)	(3.98)	(5.25)	(2.93)	(11.62)
Detached	428.6***							419.6***
	(58.20)							(57.40)
Semi-detached	175.7***							174.1***
	(33.27)							(32.85)
Terraced	Hold-out							Hold-out

Flat	91.56***							102.4***
	(3.40)							(3.76)
Tenure	20.88	32.32	32.04	24.67	20.84	27.74	-517.7*	19.17
	(1.78)	(1.19)	(0.54)	(0.85)	(1.12)	(1.54)	(-2.24)	(1.64)
Urban-rural indicator	15.67*	68.05***			-3.489	-24.15*	648.2**	14.39*
	(2.39)	(5.24)			(-0.33)	(-2.42)	(2.96)	(2.19)
Constant	1822.3***	2384.0***	2574.3***	2101.9***	1833.0***	1694.6***	2396.2***	1677.2***
	(74.20)	(38.12)	(22.86)	(25.19)	(39.63)	(53.69)	(13.86)	(46.05)
Quarterly fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Postcode fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
adj. R-sq	0.422	0.221	0.169	0.290	0.387	0.523	0.432	0.420
N	62,464	18,568	7,686	10,882	21,069	22,109	718	62,461

t statistics in parentheses * $p<0.05$, ** $p<0.01$, *** $p<0.001$

APPENDIX – data set

Field Name	Description	Data type	Descriptives
SalePrice1	1st sale price	Numeric (continuous)	n: 191,554 Mean: £144,017 Median: £125,000 Std. Deviation: £85,543 Min: £9,000 Max: £1,900,000
SaleDate1	1st sale date	Date (interval)	Range: 2 Jan 2003 – 26 Feb 2014
SalePrice2	2nd sale price	Numeric (continuous)	n: 47,158 Mean: £151,405 Median: £132,000 Std. Deviation: £81,637 Min: £10,000 Max: £1,775,000
SaleDate2	2nd sale date	Date (interval)	Range: 27 Feb 2003 – 26 Feb 2014
House Price Index	Percentage change in the Regional Land Registry House Price Index between first and second sale	Computed variable (numeric, continuous)	n = 25,189 Mean = -0.01% Std. Deviation: 13.92 Min: -22.93% Max: 42.16%
House Price Appreciation	Percentage change in price between first and second sale	Computed variable (numeric, continuous)	n = 25,189 Mean= 16.99% Std. Deviation: 37.62 Min: -85.06% Max: 316.58%
Property Type		Category (ordinal)	D (detached): 55,702 (29%) S (semi-detached): 61,153 (32%) T (terraced): 70,570 (37%) F (flat): 4,129 (2%)
Tenure	Legal interest	Category (nominal)	F (freehold): 179,802 (94%) L (leasehold): 11,751 (6%) U (unknown): 1 (-)
Beds	Number of bedrooms	Numeric (interval)	0: 53 (-) 1: 2,864 (2%) 2: 38,602 (20%) 3: 83,659 (44%) 4: 26,374 (14%) 5: 4,952 (3%) 6: 1,173 (1%) 7: 288 (-) 8+: 194 (-) Missing: 33,395 (17%)
YearBuilt	Year of construction (age categories were computed from this variable)	Date (interval)	<1900: 7,669 (4%) 1900: 15,790 (8%) 1901-1929: 15,088 (8%) 1930-1949: 10,263 (5%) 1950-1959: 8,989 (5%) 1960-1969: 9,178 (5%) 1970-1979: 10,323 (5%) 1980-1989: 6,955 (4%) 1990-1999: 11,260 (6%) 2000-2009: 18,014 (9%) 2010-2013: 8,019 (4%) 2014 - : 21 (-) Missing: 69,985 (37%)
Postcode2	Unit postcode	String	
Parea	Postcode area	Category	CF: 64,882 (34%)

		(nominal)	CH: 8,806 (5%) GL: 1 (-) HR: 279 (-) LD: 2,797 (2%) LL: 34,136 (18%) NP: 29,232 (15%) SA: 45,844 (24%) SY: 5,637 (3%)
InspectDateEPC	Date of EPC inspection (EPC)	Date (interval)	Range: 7 Feb 1988 - 31 Jan 2014
LodgeDateEPC	Date of EPC lodgement (EPC)	Date (interval)	Range: 22 Apr 2007 – 31 Jan 2014
FloorAreaEPC	Total floor area (EPC)	Numeric (continuous)	n = 147,116 Mean: 98 Median: 88 Std. Deviation: 52 Min: 2 Max: 8,412
Energy_Rating_Current	Current energy rating (EPC bands were computed from this variable)	Numeric (interval)	n = 191,553 Mean: 58 Median: 60 Std. Deviation: 15 Min: 0 Max: 111 Missing: 1
RU11IND	Rural-urban classification (2011 Census)	String	A1 Urban major conurbation: OA falls within a built-up area with a population of 10,000 or more and is assigned to the 'major conurbation' settlement category. The wider surrounding area is less sparsely populated; B1 Urban minor conurbation: OA falls within a built-up area with a population of 10,000 or more and is assigned to the 'minor conurbation' settlement category. The wider surrounding area is less sparsely populated; C1 Urban city and town: OA falls within a built-up area with a population of 10,000 or more and is assigned to the 'city and town' settlement category. The wider surrounding area is less sparsely populated; C2 Urban city and town in a sparse setting: OA falls within a built-up area with a population of 10,000 or more and is assigned to the 'city and town' settlement category. The wider surrounding area is sparsely populated; D1 Rural town and fringe: OA is assigned to the 'town and fringe' settlement category. The wider surrounding area is less sparsely populated; D2 Rural town and fringe in a sparse setting: OA is assigned to the 'town and fringe' settlement category. The wider surrounding area is sparsely populated; E1 Rural village: OA is assigned to the 'village' settlement category. The wider

			surrounding area is less sparsely populated; E2 Rural village in a sparse setting: OA is assigned to the 'village' settlement category. The wider surrounding area is sparsely populated; F1 Rural hamlet and isolated dwellings: OA is assigned to the 'hamlet and isolated dwelling' settlement category. The wider surrounding area is less sparsely populated; F2 Rural hamlet and isolated dwellings in a sparse setting: OA is assigned to the 'hamlet and isolated dwelling' settlement category. The wider surrounding area is sparsely populated.
SalePrice1psm	Sale price per square metre for transaction #1	Computed variable	n = 147,116 mean = 1,537 sd = 782 range = 49,942, min = 33 max = 49,975
SalePrice2psm	Sale price per square metre for transaction #2	Computed variable	n = 37,043 mean = 1,654 sd = 700 range = 39,601, min = 33 max = 39,634