Monograph:
Hincks, S. and Leishman, C. (2021) A statistical revaluation of council tax subjects in Wales. Report. UK Collaborative Centre for Housing Evidence (CaCHE)

UK COLLABORATIVE CENTRE FOR HOUSING EVIDENCE

A statistical revaluation of council tax subjects in Wales

Stephen Hincks (Department of Urban Studies and Planning, University of Sheffield) and
Chris Leishman (University of South Australia, and University of Sheffield)

17 February 2021

About the authors

Dr Stephen Hincks is a Reader in the Department of Urban Studies and Planning at the University of Sheffield.

Prof. Chris Leishman is Professor of Property and Housing Economics, University of South Australia and Co-Investigator of the UK Collaborative Centre for Housing Evidence.

Acknowledgements

The authors would like to extend their thanks to the Local Government Strategic Finance Division at the Welsh Government for partially funding this study and to Dr John Broomfield, Clare Blake and Dominic Evans for their advice and guidance over the course of the project.

Thanks also to the ESRC for partially funding the study through the UK Collaborative Centre for Housing Evidence (CaCHE), to Gareth Young (CaCHE), and to the Valuation Office Agency for providing access to data needed to undertake the hedonic modelling work. The usual disclaimers apply.

Contents

UK COLLABORATIVE CENTRE FOR HOUSING EVIDENCE

Executive Summary

This report documents a statistical exercise undertaken by the University of Sheffield to determine the likely impacts of a revaluation of council tax subjects in Wales. The report is positioned from a policy-perspective in relation to established and long-running debates on the design and introduction of the council tax in the UK.

An initial stage of descriptive statistical analysis is undertaken to establish broad housing market trends that have been observed in Wales since the last revaluation in 2005 based on 2003 prices. This descriptive component of the study foregrounds a more detailed and complex statistical modelling exercise in which 'hedonic modelling' is used to estimate the relationships between housing and neighbourhood attributes, and transaction prices or market values across Wales between 1995 and 2019, using HMLR, VOA and publically available geospatial datasets.

The approach employed in this study extends existing valuation approaches in three principal ways. First is in the use of a variant of the standard hedonic model (the multilevel hedonic model) that enables a system of weights or 'premia' to be estimated that are shown to vary at a small spatial scale within Wales. Second is the longer time period used to collect transaction data when compared to other studies (1995-2019), which maximises the possibility of us observing transactions for as large a sample of dwellings as possible. Finally, we match housing transactions data to a rich dataset of variables that measure neighbourhood quality and, in particular, environmental amenity that enables the modelling to capture locality as well as dwelling characteristics in the determination of price.

The simple descriptive analysis of median transaction prices observed in HMLR data reveals a number of interesting trends in housing price change over the transaction period. The impact of the Global Financial Crisis is especially notable as price appreciation accelerated prior to 2007/08 but slowed in the period following. In addition, and significantly for the focus of this study, is a notable pattern that appears to be one of stronger growth in median prices at the lower compared to the upper end of the housing market in Wales since 2003, which marks the last revaluation year.

The insights from the simple descriptive exercise provides an important rationale for the subsequent use of the hedonic framework in that the trends observed at the lower end of the market could have implications for the positioning of band limits and would almost certainly have knock-on effects in the spatial distribution of the council tax base within Wales. The multi-level hedonic model we developed, which included structural attributes of the property, neighbourhood attributes and environmental attributes has an adjusted R square 0.82, or explains 82% of variation in property transaction prices that is reflected in small scale spatial variation across Wales. Overall, the analysis confirms the value of exploring further, the positioning and spatial distribution of the council tax base within Wales since the last revaluation exercise in 2003.

Following the hedonic pricing exercise, we subsequently turn to examine the possibilities offered by a range of revaluation and council tax banding approaches and options, paying particular attention to spatial impacts. In doing so, we test alternative banding approaches and add to the existing evidence base in this area by deriving a banding profile based on a 'pure revaluation' approach to council tax banding derived using a combination of the results from the hedonic modelling and cluster analysis.

A key conclusion from this research is that statistical methods can be used to great effect, when combined with appropriate high quality data on dwelling attributes and transactions. The modelling exercises reported resulted in the development of sophisticated, yet transparent and understandable, tools that can be used in the estimation or market value. The predictive accuracy of these tools is high, and arguably compares favourably with more time-intensive, traditional approaches. More manual, traditional, approaches will probably always be needed to reflect the valuation of atypical dwelling types, sizes or locations. Yet, we would argue that the approaches piloted and reported in this study indicate that statistical methods are an effective approach that could be used to allow more frequent and systematic monitoring of current property values in the future.

Introduction

This report documents a statistical exercise undertaken by the University of Sheffield to determine the likely impacts of a revaluation of council tax subjects in Wales. The research team received funding from the Economic and Social Research Council (ESRC) and the Welsh Government to undertake this study. It has been published as an independent research output of the ESRC funded UK Housing Evidence Centre (CaCHE).

The report has been written in a brief and accessible format, reflecting that it is not intended as an academic piece of work, but as a policy-relevant study based on robust evidence. In the next section we briefly review the background to the design and introduction of the council tax, noting that this material is already well rehearsed by previous, and some very recent, studies. Subsequent sections describe the statistical methodology in some detail before moving on to examine broad housing market trends that have been observed in Wales since the last revaluation in 2005 (based on 2003 prices). This simple and descriptive analysis provides background and sets some expectations about the likely findings of the more complex statistical analysis that follows.

The main statistical method we used to estimate the current market value of all dwellings in Wales is known as 'hedonic modelling'. This is a standard form of modelling used in estimating the relationships between housing and neighbourhood attributes, and transaction prices or market values. It is used in mass appraisal contexts, and has been heavily applied in the field of housing economics to study changes in housing prices over time and space (between and within local markets and submarkets). Although hedonic models have a long history in the analysis of housing market data, methodological improvements are often possible. We extend previous studies in three principal ways:

- We use a variant of the standard hedonic model (the multilevel hedonic model) to arrive at a system of weights or 'premia' that vary at a small spatial scale within Wales, and that also vary over time. This allows us to capture differential house price movements between locations and between property types (or sizes / market segments) more effectively than standard approaches.

- We use a much longer time period to collect transaction data than other studies (1995-2019). This long time period maximises the possibility of observing transactions for as large a sample of dwellings as possible. This is important because some dwellings – particularly large, 'upper end of the market' dwellings - transact infrequently. When hedonic models are estimated using a short time period of transactions data, the risk of sample selection bias is greater. This can distort the modelling results.

- We match housing transactions data to a rich dataset of variables that measure neighbourhood quality and, in particular, environmental amenity. For example, we measure proximity to woodlands, parks, 'blue space', conservation areas, national parks, and a range of urban amenities including roads, schools, retailing and railway stations. By including such variables in the model, the explanatory power is raised and the predicted prices should more accurately represent real market values.

Previous research on council tax revaluation

The Council Tax was introduced hurriedly, as a replacement for the unpopular Community Charge (Poll Tax), in April 1993, and has been the subject of recurring waves of criticism and debate since. Criticisms partly reflect the design of the tax which, in turn, reflect the hasty nature of its design and introduction. Others have criticised its hybrid nature, as a combination of a poll tax and property tax (see Hill and Sutherland, 1991; and more recently IFS, 2020). The basic structure of the tax is also widely acknowledged to be regressive in nature (Plimmer et al, 1999; Kenway and Palmer, 1999; Leishman et al, 2014), although a combination of discounts, exemptions and the Council Tax Reduction Scheme (previously Council Tax Benefit) help to flatten the tax. A combination of the number of tax bands, the multipliers applied to them in order to arrive at tax liability, and the number of properties in each band conspire to mean that the tax burden is much higher at lower bands both in terms of proportion of property capital value, and proportion of income of resident households.

The regressive nature of the tax also gives rise to housing market instability (Muellbauer and Cameron, 2000; Muellbauer, 2005) by effectively stoking demand for more expensive properties and for second homes / holiday homes. Jones et al (2006a) argue that there is a strong equity argument for frequent revaluations of the council tax given that housing prices tend to grow more quickly in more prosperous parts of the country. In effect, failure to revalue frequently acts to increase the regressivity of the tax. Failure to carry out regular revaluations will also accelerate polarisation between high and low price growth areas because the tax liability of the former will fall over time relative to the latter, expressed as a proportion of capital value (Jones et al, 2006b). While these distortions are inequitable, a more pressing problem may be the knock-on consequences to the distribution of central (or devolved) government grant to local authorities.

The Lyons Inquiry into Local Government (Lyons, 2007) led to renewed interest in the possibility of Council Tax revaluation and reform in England. Mirrlees (2011) also spelled out robust arguments in favour of reform, but pointed out that a combination of the unpopularity of the tax and the vocal nature of losers following revaluation give rise to a general unwillingness for governments to undertake revaluations. Nevertheless, the group of experts noted "But as council tax valuations have passed the milestone of being 20 years out of date, the absurdity of the status quo becomes ever more apparent".

A number of commentators have argued for reforms to introduce change to the tax multipliers operating between bands (Palmer, 1999; Jones et al, 2006a,b; Leishman et al, 2014; IFS, 2020). Indeed, Jones et al (2006b) reported on a hypothetical (statistical) revaluation of the Council Tax in Scotland and concluded that a simple revaluation would not do much to address the regressive nature of the tax. Their study showed that households' bills for band D dwellings would drop by as much as 5.6% or increase by as much as 9.9%, reflecting differential movement in prices across the country. A combination of revaluation and rebanding would decrease bills by as much as 20%, or increase them by as much as 26%. They examined the potential impacts of adding two bands, and to increasing the tax multiplier range from the original 0.67:2 range to 0.56:2 or 0.56:4, and concluded that the multiplier has a much greater impact both on bills and on local authority finances.

Leishman et al (2014) carried out a statistical revaluation for England based on 2011 census data, and concluded that a simple revaluation would reduce the tax take by as much as 5% in some regions, while increasing the take by up to 15% in London. However, moving towards a more progressive tax, such as a flat rate national property tax on value, would have much more profound distributional impacts, such as an 8% increase in the South East, 43% increase in London, with corresponding decreases of 15-20% in the north of the country.

Of course, Wales is unique in Great Britain in the sense that it is the one part of the country that has actually undertaken a Council Tax revaluation (in 2005, based on 2003 values). Prabhakar (2016) reports that the 2003 revaluation had the effect of increasing the tax base by 5.4% on average, increasing the revenue raised in 2004-5 from £924.1M to £1,012M. A source of controversy at the time was that the number of properties that moved up a band was much higher than anticipated due in part to the required timing of statistical projections to set the new structure and house price growth over that period (438,760 or 33% of all properties, compared to an expected 25% of properties. Meanwhile, only 8% dropped by one band). The buoyant housing market conditions at the time contributed to this effect, but Prabhakar (2016) argues that the complexity of the revaluation exercise in 2003/2005 was not communicated effectively to the public, and contributed to the failed/postponed revaluations in England and Wales in 2015.

Table 1 is reproduced from VOA (2015). It clearly demonstrates the substantial decline in the number of properties in bands A and B following the revaluation, as well as the proportional rise in bands F, G and H. The revaluation exercise also introduced a new band I with a tax multiplier of 233% relative to band D (compared to 200% for band H). However, the number of dwellings in this band is comparatively small.

Table 1 Effect of the 2005 revaluation

	Number of properties in band			
Bands	1993 list	2005 list	Number of properties that reduced to this band on introduction of the 2005 list.	Number of properties that increased to this band on introduction of the 2005 list.
A	255,840	199,480	27,470	N/A
B	325,900	284,490	23,920	72,090
C	265,000	289,030	26,050	107,290
D	200,520	206,120	17,030	82,490
E	164,120	168,260	7,180	72,160
F	64,450	103,280	3,650	61,320
G	38,250	49,190	90	26,620
H	3,390	12,050	N/A	11,250
I	N/A	5,550	N/A	5,550
Total	1,317,450	1,317,450	105,380	438,760

Source: VOA, 2015

Table 2 shows the effects of the 2005 revaluation on the council tax band limits in Wales. It reveals that the lower limits of bands A and B increased by a factor of approximately 1.47, while bands C through G increased by a factors of between 1.67 and 1.86. Band H increased by a factor of only 1.35, while the 2005 lower limit of the new band (I) is approximately 1.77 the 1993 lower limit of band H. This may indicate either that higher value properties increased much more slowly than those in lower bands between 1991 and 2003, or that the revaluation may have inadvertently increased the regressivity of the tax. However, to the best of our knowledge, there is no published analysis of the consequences of this phenomenon.

Table 2 Council Tax band limits pre and post 2005 revaluation

	Property values (£)	
Bands	1993	2005
A	< 30,000	< 44,000
B	30,001 to 39,000	44,001 to 65,000
C	39,001 to 50,000	65,001 to 91,000
D	50,001 to 66,000	91,001 to 123,000
E	66,001 to 90,000	123,001 to 162,000
F	90,001 to 120,000	162,001 to 223,000
G	120,001 to 240,000	223,001 to 324,000
H	>= 240,001	324,001 to 424,000
I	N/A	>= 424,001

Study approach and methodology

This is the final report on a study commissioned by the Welsh Government, designed to provide an up-to-date statistical revaluation of Council Tax subjects in Wales. We note that there are just over 1.4M dwellings in Wales, comprising 0.990M owner occupied, 0.203M privately rented and 0.226M social rented dwellings. This compares to 1.275M dwellings as at the 2001 census (a net increase of around 125,000 dwellings). Around 4,000 dwellings were demolished between 2001 and 2017.

The central element to our methodology is the hedonic regression (or mass appraisal method), which is almost certainly the most heavily used and tested of the various alternatives that exist to estimate the market value of a very large number of dwellings within a short period of time. The method involves estimating one or more hedonic regression models in which observed transaction prices (or sometimes values) are regressed on an array of physical dwelling attributes and, sometimes, neighbourhood quality proxies. This yields a set of coefficients, which can be thought of as implied or implicit attribute prices, e.g. the value of one bedroom, the value of 1 square metre of living space, the value of a public room etc.

There are several important empirical issues to be addressed in any hedonic modelling exercise, and inappropriate decisions in the model design, specification and estimation processes can lead to mistakes (biases) in the results that can, in turn, have important consequences for the price predictions. These can be summarised (in no particular order of importance) as follows:

- Measurement error
- Omitted variable bias
- Spatial errors
- Sample selection bias

Measurement error occurs when there are errors in variables. In the context of an hedonic regression, this might include estimates of floor area, number of bedrooms, dwelling age, or distance from amenities, and so on. The effect of measurement error is to make parameter estimates inconsistent, i.e. different to their 'true' values. The problem is not mitigated by larger sample sizes, by definition. In linear models, the tendency is for parameter estimates to be inflated. In non-linear models, the direction of the bias is more difficult to determine, and this is likely to be the case in hedonic models of dwelling prices, because past research strongly suggests the existence of several non-linearities. For example, additional bedrooms or floor area are rarely found to add value at a consistent rate, but commonly at a diminishing marginal rate.

Omitted variable bias is, broadly speaking, a term which refers to heteroscedasticity in the regression error term. Heteroscedasticity occurs when the absolute size of the residual (which is captured by the error term) is not entirely independent, but varies in relation to some unmeasured phenomenon or variable. For example, if a particular influence on property values was thought or known to be important (such as the quality of a view, for example) but impossible to measure, then we would expect the residuals from an hedonic regression to have a relationship with that unmeasured effect. The consequence of heteroscedasticity is to cause bias in the standard errors. In practice, this normally means that one or more regression parameters appears to be more significant than it is, and the overall explanatory power of the model itself can appear greater than it really is.

Spatial errors are a particularly common problem in regression models of housing markets and/or values. Housing economic theory directs us to estimate one model for each spatial or sectoral submarket or group of dwellings that obey the 'law of one price' (equality in attribute prices after allowing for spatial differences arising from differential travel costs). Estimating models for larger areas or multiple sectors comprising multiple markets / submarkets results in biased coefficients, and unreliable models. This is a particularly important factor to consider when using pooled data covering many time periods. For example, if we were to assemble a dataset spanning the period 2003-2019 (approximately covering the last Welsh revaluation to the present time), fitting a single hedonic model would implicitly assume that property values have increased at the same rate over time uniformly within the country. This is unlikely to be true, as discussed in the previous section, and the practical consequence of making this assumption would be to introduce bias to the econometric estimates.

Sample selection bias is another empirical issue that routinely affects hedonic modelling exercises. Regression analyses are usually based on observed data on property transactions, but there are several reasons for supposing that the sample of transactions that occurs in a given time period is not random, but is partly determined by the cyclicality of the housing market and the economy. For example, it is reasonable to suppose that 'starter homes' are over-represented in the transactions mix early in an economic cycle, and that higher value homes are over-represented later in the cycle. Systematic differences in the composition of the transaction mix over time can lead to bias in the hedonic regression parameter estimates.

Estimation strategy
We adopted the hedonic regression approach to model property prices in Wales in order to derive a set of models that could be used to robustly predict prices as at a specified valuation date (the third quarter of 2019). We adopted more than one model to ensure that the relationships between property values, location and characteristics were adequately captured (this is explained in more detail later in the section).

The creation of an appropriate dataset for the analysis was a clear challenge during the project, and it is important to distinguish between what is ideal, what is possible, and the consequences of making compromises on this issue. The ideal situation would be that a complete stock database with a large number of physical dwelling variables including, for example, number of bedrooms and public rooms, internal area, external area, parking arrangements, property age, property type, condition, construction type, heating arrangements, bathrooms, etc were made available to the analysis. The difficulty is that no such database exists in a publicly accessible form, and compromises are therefore inevitable.

The modelling work therefore used 'price paid' data from HM Land Registry (HMLR) matched to the limited number of variables available from the Valuation Office Agency (VOA). To this, we added more variables derived from a range of sources, as described below.

- *Flood zone high risk*. Flood zone (Natural Resources Wales) is a Flood Map highlighting areas that could be affected by flooding from rivers or the sea. Flood Zone 2 is NRWs best estimate of the areas of land between Zone 3 and the extent of the flood from rivers or the sea with a 1000 to 1 chance of flooding in any year. It also includes those areas defined in Flood Zone 3. Intersections of property locations with Flood zone 2 areas was undertaken in GIS to identify units where housing units and flood zones coincided.

- *Distance to blue space*. Blue space refers to rivers streams canals or lakes. Rivers, streams and canals were extracted from the OS MasterMap Water Network Layer. Lakes were identified from the Welsh subset of an inventory of standing waters (freshwater lakes and lochs) derived from Ordnance Survey digital map data at a scale of 1:50 000 (Natural Resources Wales). Distance to blue space was calculated using Euclidean (crow-flies) measures in GIS automated using 'Near Analysis'.

- *Distance to coast*. Distance to coast was calculated for each dwelling unit using Euclidean (crow-flies) measures in GIS automated using 'Near Analysis' with a polyline of the coastline used as the feature of interest.

- *Distance to national park*. This was measured using a multi-distance buffer of 0.5, 1, 1.5 and 2km banding. A spatial join linked individual housing units to a distance band.

- *Conservation area within 250m / 500m*. Conservations area boundaries were extracted from Lle (Welsh Government).

Two buffer bands were defined at 250m and 500m. A spatial join linked individual housing units to a distance band.

- *Walk time to green space*. Green space was defined as accessible parks and open space captured through the OS Open Green Space dataset (Digimap). Walk time to green spaces was calculated using the Access points for individual green spaces and OS MasterMap Highways Network. Network distance was used to calculate walk time based on 5 minute interval bands between 0 and 60mins.

Detailed approach
The main method of analysing and predicting dwelling values used in this study is a statistical approach known as 'hedonic modelling' that is well-established in the field of applied economics, and is often used in property valuation and 'mass appraisal'. Within hedonic modelling there are various other, more detailed, methodological options. Later in the report we explain the innovations we introduce in order to reduce the possibility of bias in the forms described above. In this section we summarise the data matching, scale of attrition and approach to assembling the final database of revalued dwelling values.

The HMLR and VOA data were matched on the basis of address data. In practice, this also involved a substantial amount of manual 'data cleaning' work to ensure consistency between the key fields in these datasets – particularly in relation to street numbering, description of flat positions, and street names. Small differences in recording conventions between the two main datasets lead to a drop in the successful match rate. Ultimately, these data compilation tasks yielded a master dataset of 1,440,021 observations (all VOA records), for which there was an observed transaction price at some point during the 1995 through 2019 study period for 463,271 records.

Our use of a long study period maximised the probability of observing at least one transaction price for each dwelling in the VOA dataset. Where multiple transactions were found over time, the most recent transaction price and date were used and the others were discarded. However, we should acknowledge that for many dwellings, either no transactions occurred during the study period, or the quality of addressing fields in one dataset or the other was inadequate to allow matching.

Subsequent data cleaning phases yielded further attrition. For example, erroneously recorded postcode fields meant that lower super output area (LSOA) codes could not be matched to all records. This reduced the dataset from 463,271 to 460,383 records (this number of observations was used to estimate the first econometric model, detailed later in the report). Adding neighbourhood and environmental variables also led to some attrition and resulted in the final estimation dataset falling to 334,927 records, and the master dataset falling from 1,440,021 to 922,191 records. This can be summarised as follows:

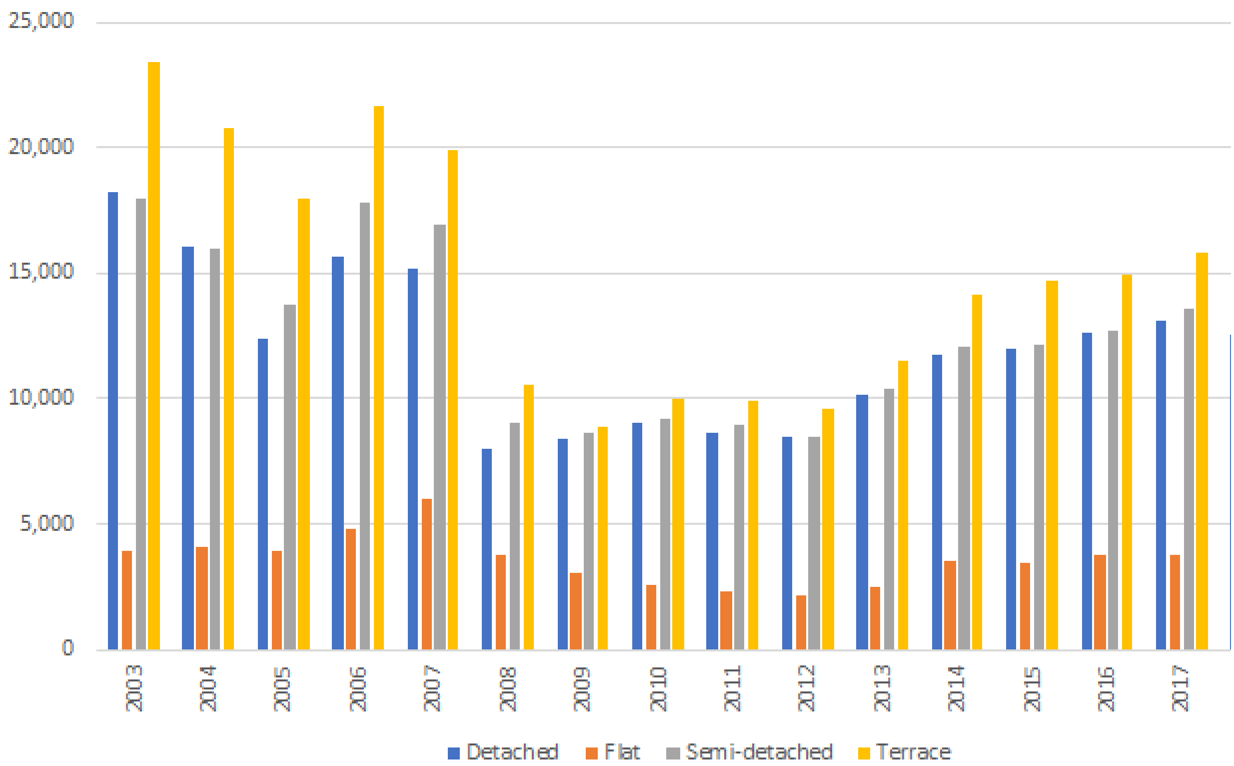
- 1,440,021 original VOA observations (dwellings)
- 460,383 with observed prices used to estimate the first model
- 334,927 with observed prices and neighbourhood/environmental variables used to estimate the second model
- 922,191 VOA records for which it was possible to predict a Q3, 2019 market value
- Scaled up to 1,440,021 using LA/2003 council tax band sampling weights.

The report provides further detail on the specification of the model in subsequent sections. The next section sets out a simple overview of housing market trends since the last revaluation in Wales in 2005 (based on 2003 market values).

Housing market trends since the last revaluation

Data on 'price paid' provided by HM Land Registry suggest that there were nearly 715,000 transactions in the period 2003-2019 in Wales. HMLR data do not include detail on property or neighbourhood level characteristics, but some analysis is possible on the basis of a broad property type variable included in the data. This allows the analysis to distinguish between detached, semi-detached, terraced and flatted dwellings. As figure 1 shows, the mix of dwellings in the transaction volume is broadly even between the three housing property types, and the proportion of flats is relatively low (around 8% on average over the period).

Figure 1 **Observed annual volume of HMLR transactions by property type**



The data also reveal a significant change in the overall volume of transactions between the most recent revaluation year (2003) and 2019, noting that only the first two quarters of 2019 were available at the time this analysis was carried out. A particularly striking drop in the transaction volume occurred with the onset of the global financial crisis (GFC), with the volume in 2008 being approximately 54% of the 2007 volume. The next four years (2009 through 2012) show a stable volume of transactions, with the volume beginning to rise noticeably again in 2013. There are then five more years of growth in the volume, with the first two observed quarters of 2019 continuing this trend.

The transactions volume is an important housing market indicator in the sense that it is often found to be a leading indicator of house price appreciation. The patterns observed in the data reflect trends that are often observed in the UK's regions, outwith London and the South East of England. Periods of heightened housing market activity and price appreciation are typically followed by periods of plateau or stability rather than periods of decline or a crash in housing prices. The early 1990s recession was an exception to this, but still saw a much more gradual adjustment in Scotland, Wales, Northern Ireland and Northern England than London and the South East.

Figure 2 summarises trends in median prices broken down by HMLR dwelling type. What is immediately apparent is the pronounced increase in prices during the period 2003-2007, followed by short adjustment around the GFC, and then a long period of stability. Median prices increased by just under 50% for detached and flatted dwellings, or 62% for semi-detached and 75% for terraced dwellings between 2003 and 2007. The price adjustment following the GFC occurred, approximately, between 2007 and 2011, with the median price falling by around 10% for all the housing types, or 14% for flats. During the next eight years, to the mid-point of 2019, prices increased slowly and reached around 17% to 20% of their 2011 level by then (for houses), or 7.4% for flats. This suggests that the housing market recovered quite slowly in the post 2011 period compared to the observed period prior to the GFC, and that the market for flats did not recover particularly strongly.

Figure 2 **Median prices for HMLR dwelling types**

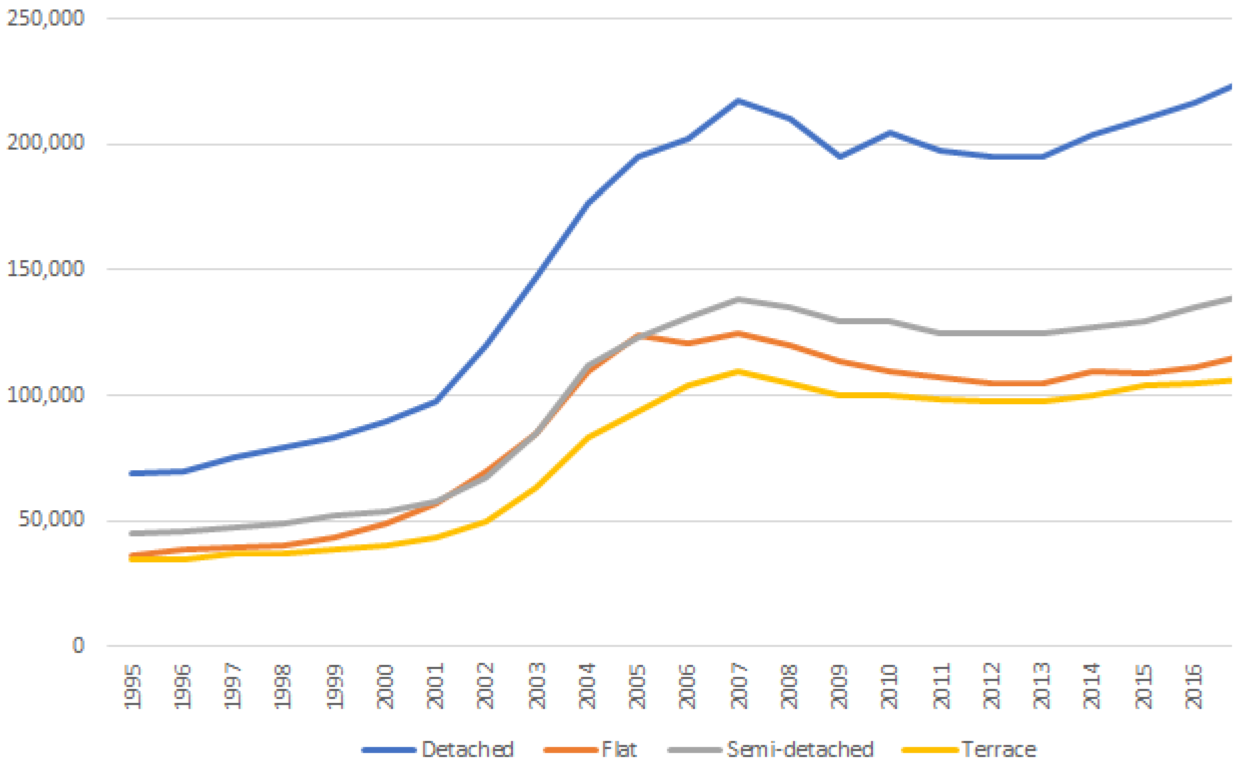
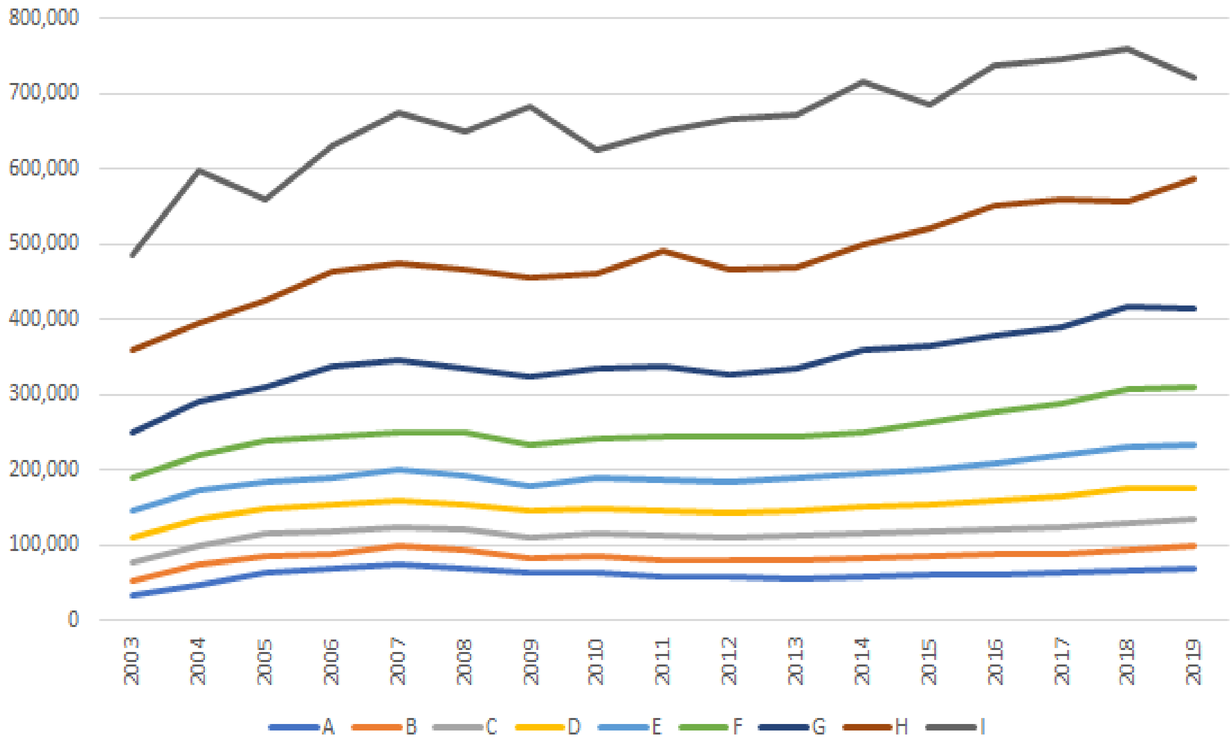


Figure 3 breaks down the analysis by 2003 council tax band, as observed in the VOA dataset and matched to HMLR transactions. The trends show that median nominal prices have increased significantly over the study period, despite the downturn and period of stagnation in and following the GFC. However, comparing median prices in 2019 relative to 2003 reveals that price growth may have been stronger in the lower than upper ends of the market across this period. For example, the median price increased from £34,000 to £69,000 (a ratio of 2.03) for band A dwellings, compared to an increase from £485,000 to £720,000 for band I (a ratio of 1.48). The ratios for bands B and C are 1.89 and 1.73, respectively. For other bands, the ratios are approximately 1.60.

Figure 3 **Median prices for 2003 council tax bands**



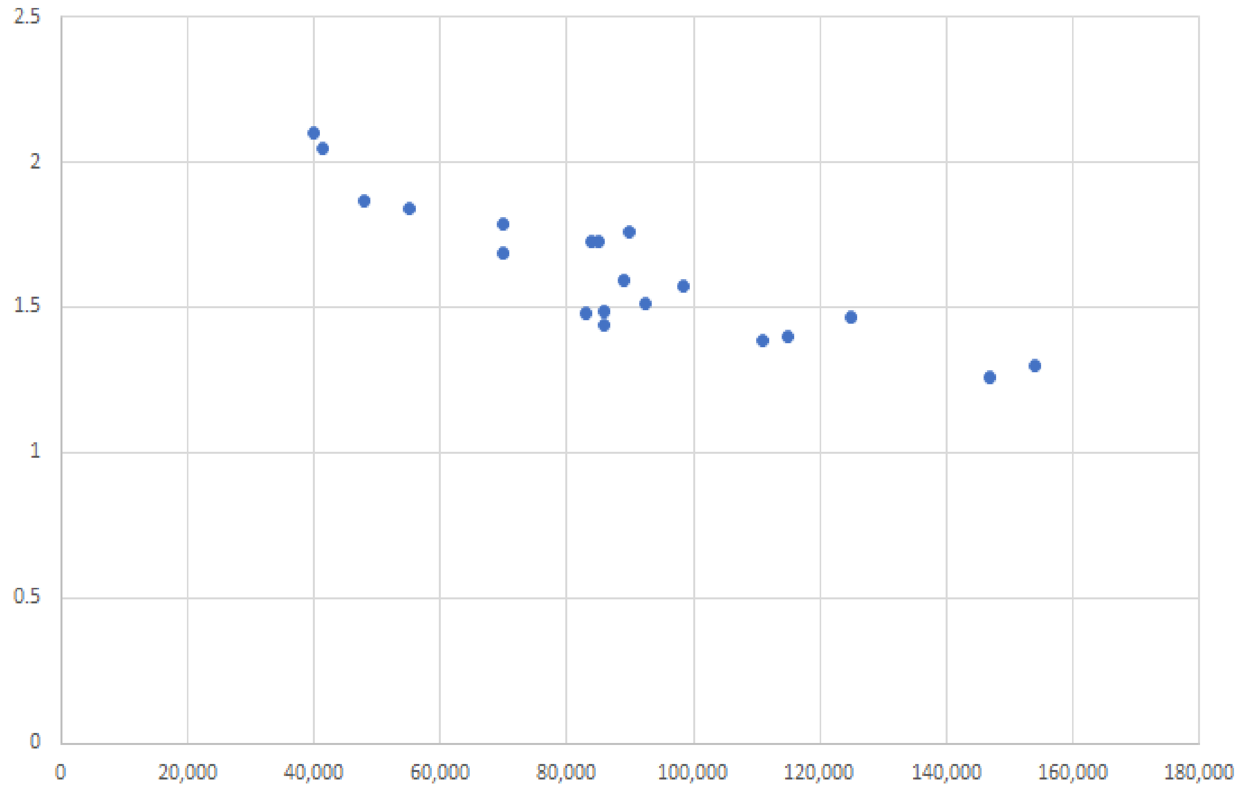
It also seems probable that there have been spatial variations in housing price change within Wales since the last revaluation in 2003, and potentially since the GFC. Table 1 shows considerable variation in housing price levels between local authority areas, with Cardiff, Ceredigion and Monmouthshire having significantly higher prices than other parts of the country. Interestingly, these local authorities had a more modest increase in median prices between 2003 and 2007 (the peak of the last housing market cycle, before the GFC). The median price rose by between 30% and 45% in these areas. Meanwhile, in Blaenau Gwent, Merthyr Tydil and Neath Port Talbot – areas with 2003 median prices in the region 40,000 to 55,000 – the median price increased by between 90% and 110%. In general, there is a strong association between (low) 2003 median price levels and high rates of increase in the median price between 2003 and 2007. This is demonstrated in figure 4 (the correlation is -0.89).

Table 1 Median prices, and medians relative to 2003 for local authority areas

Local Authority	2003	2007	Ratio	2009	Ratio	2019	Ratio
Blaenau Gwent	40,000	84,000	2.10	70,000	1.75	82,000	2.05
Bridgend	83,000	122,850	1.48	120,000	1.45	140,000	1.69
Caerphilly	69,950	118,000	1.69	103,000	1.47	128,000	1.83
Cardiff	132,000	172,000	1.30	170,000	1.29	219,000	1.66
Carmarthenshire	70,000	125,000	1.79	115,000	1.64	122,500	1.75
Ceredigion	127,000	184,000	1.45	165,000	1.30	177,000	1.39
Conwy	111,500	154,250	1.38	140,000	1.26	161,000	1.44
Denbighshire	86,250	128,500	1.49	123,250	1.43	138,250	1.60
Flintshire	95,000	143,000	1.51	135,000	1.42	153,250	1.61
Gwynedd	84,000	145,000	1.73	140,000	1.67	151,500	1.80
Isle of Anglesey	85,000	147,500	1.74	145,000	1.71	160,000	1.88
Merthyr Tydfil	41,500	85,000	2.05	74,500	1.80	85,000	2.05
Monmouthshire	145,000	200,000	1.38	175,000	1.21	240,000	1.66
Neath Port Talbot	55,000	104,500	1.90	92,500	1.68	108,250	1.97
Newport	94,000	140,000	1.49	129,950	1.38	163,250	1.74
Pembrokeshire	98,500	155,000	1.57	156,000	1.58	150,000	1.52
Powys	110,000	170,000	1.55	155,000	1.41	174,000	1.58
Rhondda Cynon Taf	48,000	89,000	1.85	80,000	1.67	91,500	1.91
Swansea	82,000	131,950	1.61	134,000	1.63	140,000	1.71
Torfaen	79,500	120,000	1.51	110,750	1.39	133,000	1.67
Vale of Glamorgan	115,000	163,000	1.42	170,000	1.48	197,750	1.72
Wrexham	89,000	141,000	1.58	135,000	1.52	148,000	1.66

UK COLLABORATIVE CENTRE FOR HOUSING EVIDENCE

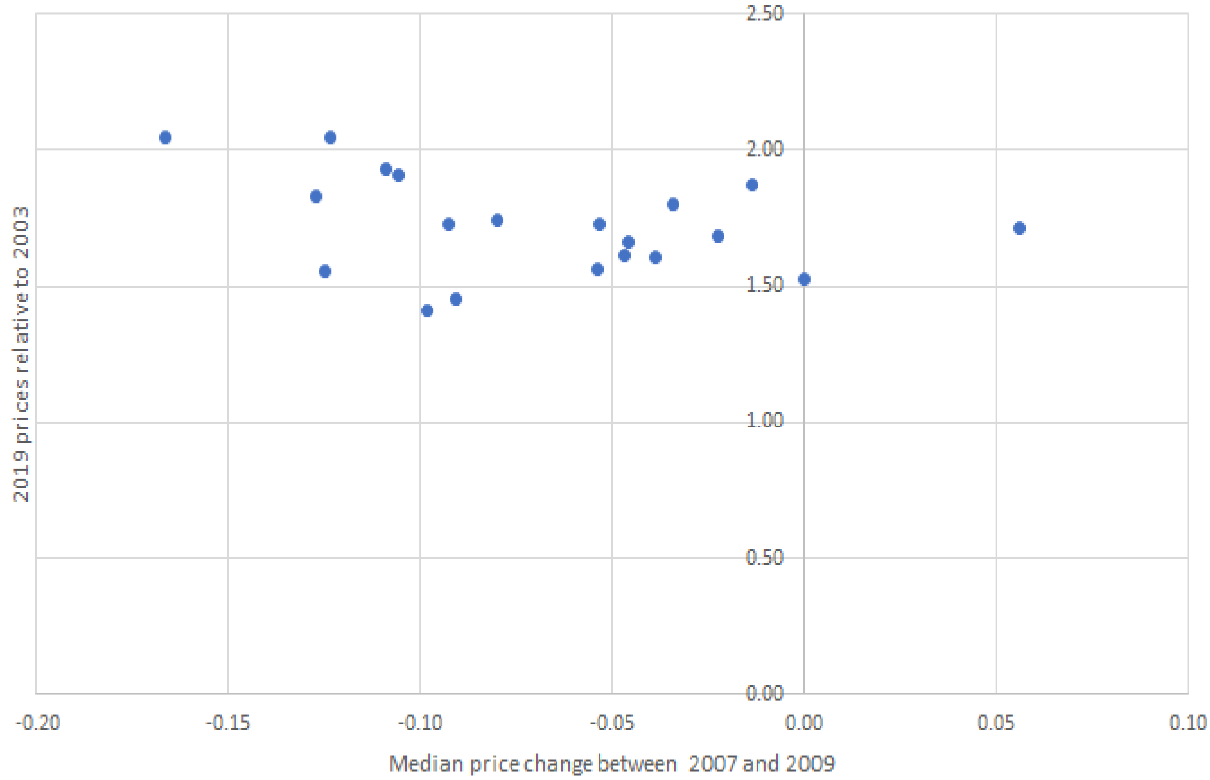
Figure 4 2003 to 2007 price growth relative to 2003 prices



The spatial pattern of growth after the GFC is distinctly different to that observed before – with some exceptions. For example, the three local authorities just mentioned, with low median prices in 2003 and high growth between 2003 and 2007, also experienced high growth in the post GFC period. However, aside from these examples the growth in median prices elsewhere in Wales was more uniform between 2008 and 2019 in most local authorities with the result that we should expect the distribution of housing prices to remain largely unchanged for much of the country.

Other exceptions relate to the local authorities with relatively high median prices in 2003 / 2007. Cardiff, Ceredigion, Conwy and Monmouthshire had median prices in 2003 of 132,000, 127,000, 111,500 and 145,000 respectively, and experienced the lowest median price growth rates between 2003 and 2007, and between 2009 and 2019. Indeed the correlation between the 2019 to 2003 ratio of prices and the 2009 through 2019 price ratios is 0.17. This is also demonstrated in figure 5 which reveals that there is almost no visibly discernible relationship between areas which experienced a significant price correction between 2007 and 2009, and the rate of median price change afterwards.

Figure 5 Relationship between GFC price change and subsequent price growth



In conclusion, a simple descriptive analysis of median transaction prices observed in HMLR data reveals a number of interesting trends that are worthy of further consideration in the more detailed subsequent analysis. The most notable pattern appears to be one of stronger growth in median prices at the lower than the upper end of the market since 2003 – the last revaluation year. If confirmed by the more detailed hedonic analysis that follows, this has implications for the positioning of band limits, and will almost certainly have a knock-on to the spatial distribution of the council tax base within Wales.

Results from the main modelling phase

The data matching and cleaning operations leading to the main analysis are detailed earlier in the report. In this section we set out the estimation strategy in more detail before presenting the results.

As noted earlier, the estimation strategy involves two distinct models. The first is a multi-level model following the approach set out by Leishman (2009). The multi-level hedonic model provides considerable flexibility when modelling prices across a large spatial unit such as a metropolitan area or regional market. Sub-market theory informs us that dwelling attribute prices may diverge over space (or time) owing to inefficiencies in markets, heterogenous consumer preferences, spatial clustering of close substitutes, and differential levels of housing supply within spatial housing markets. In other words, the implicit or hedonic price of an attribute such as a garden, a home office or a fourth bedroom, may not be constant with respect to the prices of all other attributes. If the relativities change over time, or are different between the submarkets or neighbourhoods within a housing market, then the modelling strategy needs to be flexible enough to capture these dynamics. Failure to do so introduces bias to the fitted model, and can give rise to misleading results.

The multi-level hedonic model deals with this problem, to a certain extent. Dwelling attribute prices are assumed to have a common component which is constant over space within the housing market, and a heterogenous component unique to smaller units of geography. At the simplest level this can be specified as a market-level model, but one which permits a 'premium' which is specific to the observed smaller scale spatial unit (a neighbourhood or LSOA level price adjustment, in effect). Formally, this is specified as follows:

$$P_{ij} = \alpha + \Sigma \beta_k X_{kij} + (\varepsilon_i + u_{ij}) \qquad (1)$$

The specification can be refined further if there is reason to suppose that the hedonic prices of other dwelling attributes may also vary over space, as shown in equation 2:

$$P_{ij} = \alpha + \Sigma \beta_k X_{ki} + (\varepsilon_i + u_{0j} + u_{kj} X_{ki}) \qquad (2)$$

Where,

P_i Transaction price of the *i*th dwelling

P_{ij} Transaction price of the *i*th dwelling in spatial unit *j*

X_{ki} Represents a vector of physical and neighbourhood housing attributes

ε_i The standard OLS error term or residual

u_{ij} Random intercept for the *j*th spatial unit

u_{kj} Random coefficient for the kth attribute in the *j*th spatial unit

Leishman (2009) shows that the multi-level hedonic model is versatile with respect to capturing the changing relativities of hedonic prices within a metropolitan housing system, with modelling based on individual or successive annual cross-sections (i.e. building a model with a single year of observed transaction data). An additional complication arises when the hedonic model covers numerous time periods because the relativities between hedonic prices may vary over time as well as space. For example, the importance of one additional bedroom versus an additional 20 square metres of internal space may be different in a high density urban neighbourhood than in a lower density suburban neighbourhood, but may also change over time as demand and housing market pressures change.

One theoretical solution would be to specify a multi-level model in which hedonic prices may vary over time as well as space. However, this approach is inefficient and rarely leads to stable estimation results (such heavily specified models are often referred to as 'over-fitted'). On the other hand, ignoring the issue completely would increase the risk of bias, and hence misleading results, arising from the misspecification. We therefore designed a compromise approach:

- Step 1 - Construct a system of MSOA dummy variables interacted with time dummies; and a set of LSOA dummies;
- Step 2 - Estimate the multi-level hedonic model, saving the estimated random effects;
- Step 3 - Estimate the OLS hedonic model, using random effects variables as predictors.

To explain, the dummy variables described in step 1 are intended to capture several different sources of variation in hedonic prices. Simply including a set of time dummy variables is equivalent to assuming that there is a single price index for Wales over time that applies to all dwellings and locations equally. By including interactions of MSOA/time dummies, the assumption is of a different price index for each MSOA over time. This allows the model to acknowledge that an MSOA's price index can diverge from the Wales price index.

The choice of MSOA as the appropriate unit of geography is a pragmatic one in the sense that the MSOA is small enough to be considered a proxy of neighbourhood, but large enough to ensure that there are sufficient transactions in each observed transaction year to allow the model estimation. Finally, including the LSOA dummies allows the model to reflect that there may be unexplainable (or, at least, difficult to explain) variations in housing prices between neighbourhoods even after accounting for all observable differences in dwelling and neighbourhood variables.

The multi-level model described under step 1 otherwise has a simple specification that includes internal floor area, number of rooms, number of bedrooms, and property type. The estimation results are more difficult to interpret than simple OLS results, but it is important to emphasise that the main purpose of this step is to generate the two 'random effects' variables to be included in the subsequent OLS model. As table 2 shows, the coefficients for dwelling area, total rooms, bedrooms and number of bathrooms are statistically significant at 1%. Other factors held constant, property value increases by approximately 2.7% for each additional 10 square metres of internal area. Each additional room adds 2.87%, each bedroom 4% and each additional bathroom 19.55%. While these results are plausible overall, the main function of the multi-level model is to provide sets of random effects variables that measure as much constant quality spatial and temporal variation in prices as is feasible. This approach allows the main modelling phase, using a more standard OLS hedonic approach, to focus on physical and environmental attributes.

Given that the parameters are estimated using a maximum likelihood method, rather than ordinary least squares (OLS), the random effects cannot be observed directly, but are generated post estimation. The parameters of the main variables (area, rooms, bedrooms and bathrooms) are estimated directly as fixed effects. The estimation output shown in table 2 provides a range for the estimates of the two random effects variables, rather than coefficients. For example, the MSOA/YEAR dummy variables have estimated elasticities of -1.169 through 0.790. This suggests that dwelling prices may be up to 117% lower than, or as much as 79% of values in the base year (2019) and an arbitrarily chosen base MSOA.

The LSOA dummies have elasticities ranging from -0.938 to +0.934, which suggests that dwellings may have a discount of up to 93.8% or a premium of up to 93.4%, holding all other factors constant, depending on which LSOA they are located it.

The likelihood ratio test (LR statistic) is significant at 1%. In other words, we can reject the hypothesis that the random effects variables are equal to zero.

Table 2 Summary of multi-level model results

Variable	Coefficient (or range)	Significance
Constant	10.7328	***
Internal area	0.0027	***
Rooms	0.0287	***
Bedrooms	0.0401	***
Bathrooms	0.1955	***
Fitted random effects		

MSOA intercept	-0.682 through 0.580	
MSOA/YEAR	-1.169 through 0.790	
LSOA intercept	-0.938 through 0.934	
LR test	5.0e+05	***
N	460,383	
Number of MSOA	411	
Number of MSOA/YEAR	3,691	
Number of LSOA	1,907	

We now proceed to the results of the second stage model. As noted earlier, this is a more standard hedonic model estimated using OLS. The random effects variables saved after the multi-level estimation are included as predictors or controls on spatial and temporal variation in property and neighbourhood variables. In addition to controlling for the introduction of bias to the parameter estimates, this approach has the advantage of allowing us to observe the impact of a set of environmental variables on property values, after the most dominant neighbourhood effects have been taken into account. To aid interpretation and discussion of the results, the coefficients are set out in three tables (3, 4, and 5, below). However, it is important to note that these tables refer to a single model estimation, and that the coefficients have been grouped into three tables as a convenience and to improve readability. Table 3 also includes statistics relating to the overall empirical performance of the model.

Table 3 OLS model results – physical attributes

Variable	Coefficient	Significance
Constant	10.7546	***
Area	0.0055	***
Area square	-4.58E-06	***
Area cubed	4.59E-10	***
Rooms	0.0302	***
Rooms squared	-1.00E-03	***
Rooms squared	8.91E-06	***
Bedrooms	0.0211	***
Bedrooms squared	-1.00E-03	***
Bedrooms cubed	1.99E-05	***
Bathrooms	0.1566	***
Bathrooms squared	-0.0541	***
Bathrooms cubed	6.00E-03	***
Property type (9)	-12.0% to +31.2%	All ***
Property age (7)	-12.2% to +17.6%	7 at *** and 2 at **
Parking (5)	-1.4% to +11.4%	4 at *** and 1 at **
Quarterly dummies (3)	-5.6% to +0.06%	All ***
Year dummies (24)	-33.5% to +1.2%	All ***

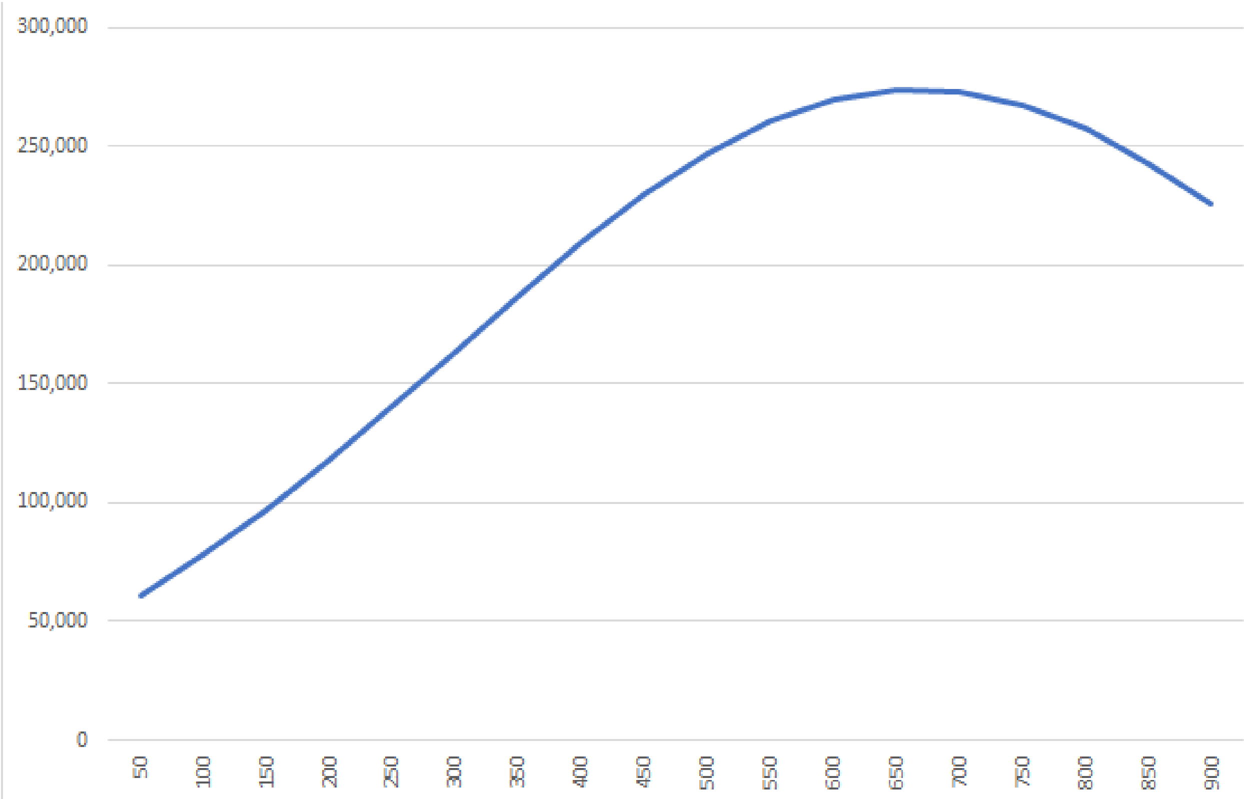
MSOA intercept	-0.022	***
MSOA / pooled year premium	0.850	***
LSOA intercept	0.689	***
Neighbourhood variables	(see next table)	
Adjusted R Square	0.822	
Std. Error of the Estimate	0.287	
F statistic	15976.833	***
N	334,927	

In choosing the functional form of the OLS model, a number of exploratory curve fitting estimations were carried out before running the full model. These results are not shown here for the sake of brevity, and because the results are uncontroversial. The exploratory estimations showed that, for most continuous variables, quadratic and cubic functional forms generated higher explanatory power than simple linear or log-linear alternatives. For this reason, dwelling area, rooms, bedrooms and bathrooms are entered as untransformed variables rather than being expressed in natural logarithms. In addition, the squared and cubed values of these variables are entered as additional variables.

The cubic specification captures several evident non-linearities. For example, area, rooms, bedrooms and bathrooms all add value, but the coefficients on the squared values of these variables are negative, suggesting that these attributes add value at a diminishing rate. However, the cubed values of the variables have positive, statistically significant coefficients. This suggests that dwellings with very large values for these variables are associated with a premium in terms of their value.

The cubic functional form can be described as an s-curve in the sense that there are two turning points on the relationship between a variable (such as area) and property value. The combined effects of the linear, quadratic and cubic influences of area on property value are summarised in figure 6. The figure is merely illustrative of the relationship between this one particular variable and price given that the predicted prices are produced by setting all other attributes to zero. Nevertheless, the demonstration is effective in showing that the impact of additional area on price increases above a certain threshold (around 150 square metres), and levels off again after another threshold (around 350 square metres).

Figure 6 **The modelled relationship between dwelling area and price**



The estimation results show that the cubic functional form holds for all four of the physical attribute continuous variables (area, rooms, bedrooms and bathrooms). There are ten property type variables altogether, meaning that nine dummy variables are entered into the model. All of these are statistically significant at 1%, and they show that there is a discount to premium range of -12% to +31%, holding all other factors constant. Property age band accounts for a discount to premium range of -12.2% to +17.6%. Parking arrangement variables are also statistically significant. Overall, the model has an adjusted R square 0.82, or explains 82% of variation in property transaction prices.

As discussed earlier, much of the spatial and temporal variation in property values, holding physical attributes constant, is captured through a system of random effects and time dummy variables. To summarise:

- Time dummy variables (year and quarterly dummies) capture the time trends in property values that broadly apply across Wales as a whole during the study period.

- LSOA random effects capture the discounts or premia that are detectable across the entire study period, and which cannot be explained with reference to physical or environmental attributes (i.e. holding these factors constant).

- MSOA random effects capture a higher level (MSOA) spatial discount or premium, but this is permitted to change over time.

The results in table 3 show that all three of these groups of spatial/temporal effects are statistically significant. The large volume of results precludes any effective analysis, but some insights can be had by mapping selected outputs. For example, figures 7 and 8 show the spatial distribution of LSOA premia (which are fixed over the study period) and the MSOA premia for 2019.

Figure 7 Spatial price premia measured at LSOA level

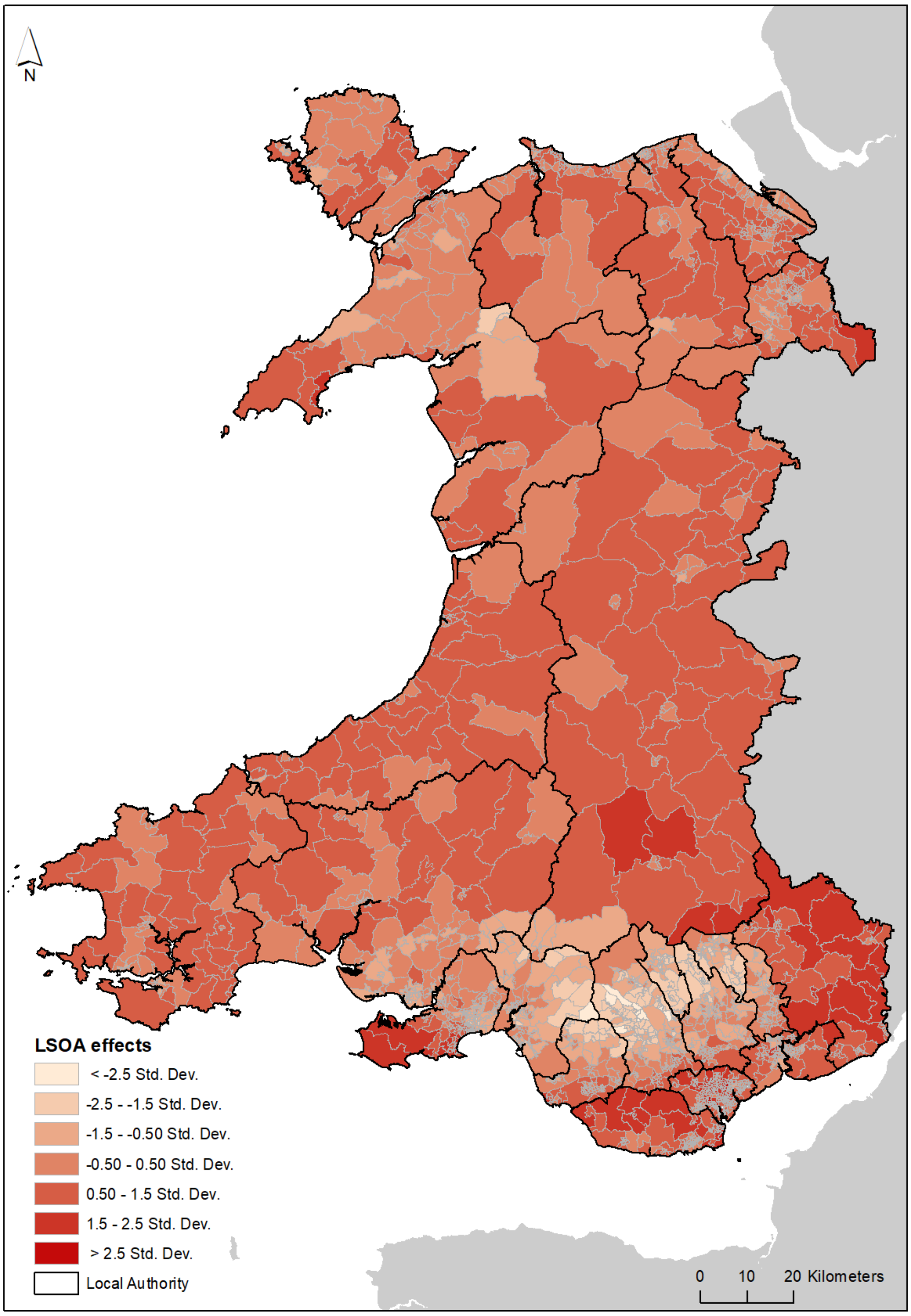


Figure 8 Spatial price premia in 2019 measured at MSOA level

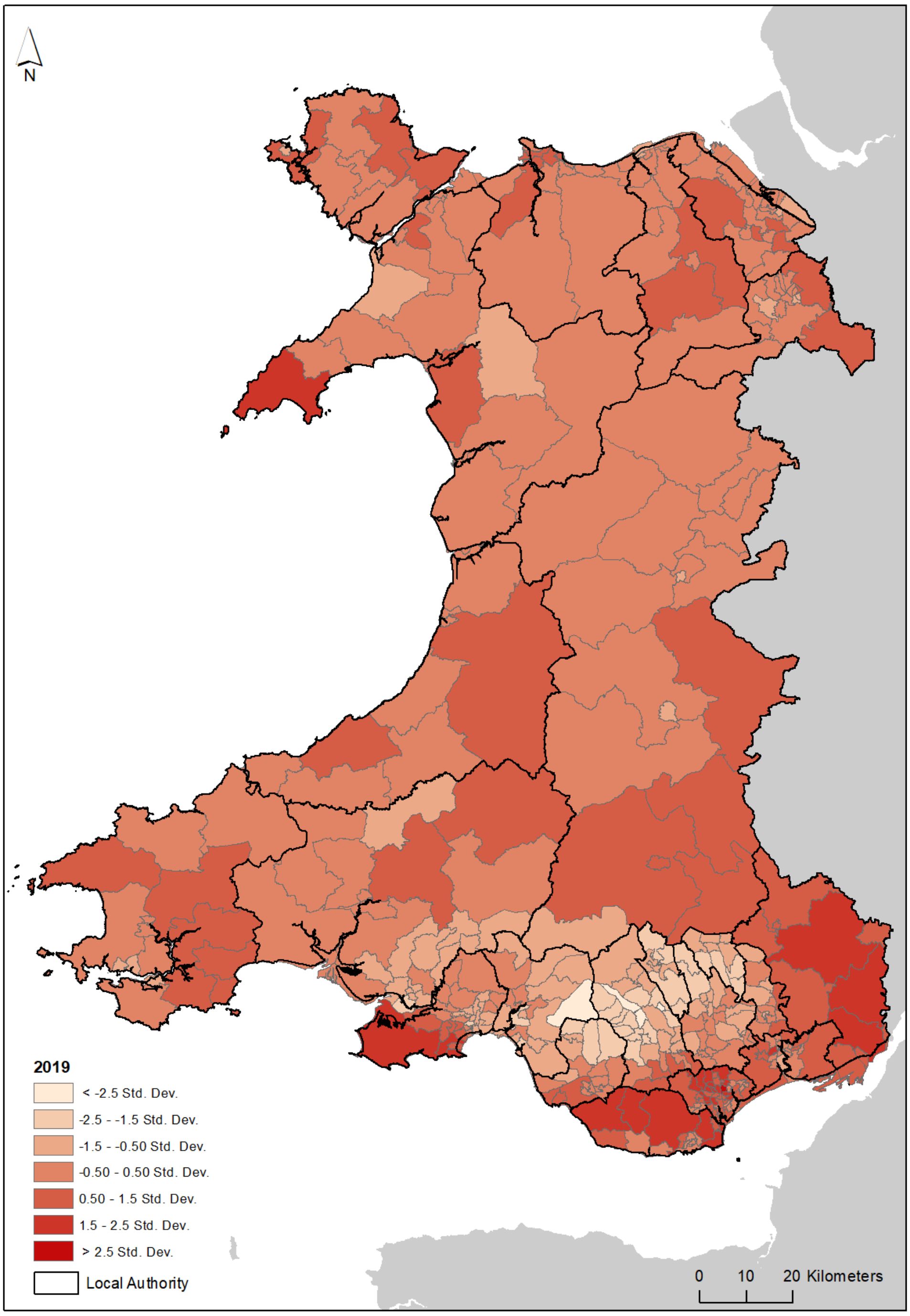


Table 4 summarises the neighbourhood variable coefficients. The final selection of variables reflects the outcomes of a number of earlier estimations designed to reveal the likely functional forms of the relationships between these variables and transaction price, and to test which of a larger set of neighbourhood variables are actually statistically significant. For example, the dataset also included information on distance and drive time to a hospital, but this variable proved to be statistically insignificant. Similarly, we found no stable relationship between distance to primary schools and property values. This does not necessarily mean that there is no relationship in reality, but may suggest that the relationship is very complex, or depends partly on other variables with an intervening effect.

As noted above, early estimations suggested a complex functional form or relationship might exist, and so distance variables raised to the powers of 2, and 3, were entered into the equation. In other cases, the relationships were simpler. For example, proximity to a railway station is associated with higher transaction prices, but the effect of distance to a supermarket is more complex. This relationship follows a quadratic functional form in which value falls with distance, but at a decreasing rate. Proximity to a secondary school, and to a main road, both follow the cubic functional form described earlier. In both cases, property values increase with distance, but at a falling rate. We found evidence that proximity to a GP surgery (within 20 minutes) is a positive influence on values, but that being within 30 minutes is a slight negative. In the final model, only the second of these is statistically significant. Although the coefficient sign seems counter-intuitive, we retain the variable on the basis that there is no compelling empirical reason to exclude it.

Table 4 OLS model results – neighbourhood attributes

Variable	Coefficient	Significance
Walk time station	-0.0014	***
Walk time supermarket	-0.0037	***
Walk time supermarket ^2	7.46E-05	***
Walk time sec. school	0.0047	***
Walk time sec. school ^2	-1.78E-04	***
Walk time sec. school ^3	1.76E-06	***
GP surgery within 30 minutes	-0.0175	***
Walk time main road	8.15E-06	***
Walk time main road ^2	-2.22E-09	***
Walk time main road ^3	1.07E-13	***

NB: ^2 and ^3 indicate increases in distance thresholds

Finally, we turn to consider the influence of environmental variables on transaction prices, as summarised in table 5. We find that location within a high risk flood zone reduces value by around 1.4%, whereas proximity to a conservation area increases value by between 2.4 and nearly 5%, depending on whether the 500m or 250m proximity measure is used. Four other proximity measures follow a similar cubic functional form explored earlier in relation to physical attributes. These are proximity to: the coast, woodland, a national park, and green space. Of course, it is possible that some dwellings will have high values for more than one of these amenity variables. The modelling approach implicitly assumes that there are no interaction effects between these. The results show that property values fall with distance from these amenities, but at an initially decreasing rate. But dwellings that are very remote from these have an additional negative premium. We also experimented with a number of other amenity variables that proved either to be insignificant, or whose coefficients were unstable. Air pollution, broadband speeds and orientation towards the sun are examples of these. Although these variables were significant in some circumstances, they ultimately proved too unstable to justify including in the final estimation.

One surprising result relates to variables measuring distance to blue space. These show unexpected signs suggesting that values increase with distance, but at a decreasing rate. Although these are difficult to explain, in the absence of a compelling empirical reason we decided to retain the variables in the final estimation.

Table 5 OLS model results – environmental attributes

Variable	Coefficient	Significance
Flood zone high risk	-0.0144	***
Distance to blue space	1.83E-05	***
Distance^2 to blue space	-2.13E-08	***
Distance^3 to blue space	5.24E-12	***
Distance to coast	-2.91E-06	***
Distance^2 to coast	1.57E-10	***
Distance^3 to coast	-2.23E-15	***
Distance to woodland	-8.80E-05	***
Distance^2 to woodland	8.51E-08	***
Distance^3 to woodland	-2.33E-11	*
Distance to national park	-1.33E-06	***
Distance^2 to national park	1.25E-10	***
Distance^3 to national park	-2.36E-15	***
Conservation area within 250m	0.0246	***
Conservation area within 500m	0.024	***
Walktime to green space	-0.0016	***
Walktime^2 to green space	3.50E-04	***
Walktime ^3 to green space	-5.74E-06	***

NB: ^2 and ^3 indicate increases in distance thresholds

Likely impacts of taxation reform

In this section we examine the implications of a range of revaluation and council tax banding approaches. We focus on the spatial impacts, i.e. the effect on the council tax base by local authority, and the distributive impacts upon households (the number and proportion of dwellings moving up or down bands, by local authority). We begin by setting new band upper value limits that would place the same proportions of dwellings in each of the current council tax bands as compared to pre-revaluation (we describe this as Option 1). The results are shown in table 6. We also include the pure revaluation band limits suggested by IFS (Adam et al, 2020) in their recent study, for the purpose of comparison.

Table 6 Likely new band limits following a pure revaluation

Band	Modelled new upper limits	IFS study limits for reference
A	92,000	78,710
B	125,000	112,490
C	163,000	155,490
D	207,000	203,970
E	272,000	275,840
F	364,000	380,670
G	532,000	553,850
H	740,000	741,180
I	-	-

With the possible exceptions of bands A and B, the band limits we propose are very close to the values suggested by IFS (Adam et al, 2020). Our proposed upper limit for band A is £92,000 or approximately 17% higher than that proposed by the IFS study, and our upper limit for band B is £125,000 or approximately 11% higher. Differences between upper limits for the other bands are small enough that they can safely be disregarded.

There are numerous possible reasons for the divergence between our estimated upper band limits for bands A and B and those proposed by IFS (Adam et al, 2020). Although both studies have used hedonic price modelling, there are methodological differences in the detailed approach. The IFS study drew on a model that involved interacting physical and neighbourhood variables with x,y co-ordinates. The modelling approach reported here used a system of spatial and temporal random effects to yield premia, and change in premia over time, at small spatial scales. There are also differences in data. In this study we opted to include all HMLR transactions over the period 1995-2019, whereas the Adam et al (2020) study focused on a shorter time period (2010-2018).

Table 7 presents a simple analysis of median prices, broken down by price band. By focusing on 2003, 2007 and 2019, the analysis reveals the very significant price growth that occurred between Wales' last revaluation year (2003) and the peak of the housing market prior to the GFC. The analysis also shows that price growth was generally higher in the lower council tax bands, and particularly so for bands A and B. For example, the median increased by 21.5% per annum in that four year period. When we extend the time frame to cover the period 2003-2019, the annualised rates of price change are considerably smaller, emphasising the dramatic effect of the pre-GFC housing boom on housing prices. We can easily see that the annualised rates of price change are much smaller when measured over this longer time period, but that there is a persistent pattern involving slightly higher levels of growth in lower than upper bands. Over a relatively long time period, the small annualised differences in price change compound.

Table 7 Median prices and annualised price change by council tax band

Band	2003 median	2007 median	Annualised change	2019 median	Annualised change
A	34,400	75,000	21.5%	70,000	4.5%
B	53,000	99,973	17.2%	100,000	4.0%
C	78,000	125,000	12.5%	135,000	3.5%
D	109,000	160,000	10.1%	175,000	3.0%
E	145,000	199,000	8.2%	232,500	3.0%
F	190,000	250,000	7.1%	310,000	3.1%
G	250,000	354,000	9.1%	415,000	3.2%
H	365,000	490,000	7.6%	598,000	3.1%
I	495,000	705,000	9.2%	775,000	2.8%

Table 8 expresses the proposed band limits under pure revaluation as a proportion of the respective 2005 band limits. This emphasises that the limits proposed by this study are relatively higher than those recently proposed by Adam et al (2020). It is worth noting that differential price growth of 1.5% per annum (for band A relative to bands D/E, from table 7 above), amounts to a differential growth of nearly 27% over a 16 year period. We would therefore argue that the limits we propose will more effectively capture the stronger price appreciation at the lower end of the market in Wales since the last revaluation.

Table 8 Pure revaluation upper band limits as a proportion of 2005 limits

Band	Adam et al (2020) limit as % of 2005 limit	Proposed new limits as % of 2005 limits
A	79%	109%
B	73%	92%
C	71%	79%
D	66%	68%
E	70%	68%
F	71%	63%
G	71%	64%
H	75%	75%
I	-	-

The analysis reported in this study also examined alternative approaches to setting new council tax band limits. Our 'pure revaluation' approach was intended to deliver the same proportion of dwellings in each council tax band as observed before revaluation, i.e. the 2005 revaluation outcome (option 1). We designed band limits that would deliver the same distribution as prior to the 2005 revaluation (the distribution of dwellings by band that existed under the 1993 valuation list). This is labelled 'option 2'. Option 3 follows the VOA methodology used in the 2005 revaluation, i.e. the upper band limits are set to the mid-point between observed median sale prices for each band. Option 4 sets band limits that would result in an equal proportion of dwellings belonging to each of the nine council tax bands. Finally, we used a data reduction approach (two step cluster analysis) to automatically determine new upper band limits based on natural breaks or clustering around means in the data (option 5). The results of these alternative banding approaches are summarised in table 9.

Table 9 Comparison of band limits under different banding approaches

Band	To replicate 2005 list (option 1)	To replicate 1993 list (option 2)	2005 revaluation method (option 3)	Equal proportion in each band (option 4)	Cluster analysis results (option 5)
A	92,000	99,000	87,000	92,000	75,300
B	125,000	136,000	119,000	105,000	101,400
C	163,000	176,000	159,000	121,000	134,000
D	207,000	227,000	210,000	137,000	171,400
E	272,000	318,000	277,000	157,000	213,300
F	364,000	416,000	371,000	184,000	265,800
G	532,000	864,000	511,000	221,000	352,600
H	740,000	-	673,000	290,000	525,500
I	-	-	-	-	-

The potential impacts of the alternative banding approaches cannot be readily interpreted from the table. To explore this further, we examine the number of dwellings moving up and down bands in general, and between local authorities. We also estimate the consequences of different banding approaches on the size of the council tax base, and the distribution of tax subjects between local authorities. However, before moving to this it is worth noting some points from table 9. First, we carried out a comparison of upper band limits that would deliver the same distribution as in the 2005 valuation list, and the limits that would be obtained if the same methodology were followed as per the 2005 valuation itself. In table 9 these results are referred to as 'to replicate 2005 list' and '2005 revaluation method', respectively (options 1 and 3). This comparison is of interest primarily because the 2005 revaluation method is thought to have exacerbated the movement of dwellings up one or more council tax bands (see Prabhakar, 2016, for example). The 2005 revaluation method involves setting the upper band limit at the half-way point between the median observed sales prices for each council tax band.

What we found is that the upper band limits derived under the two approaches are very similar indeed for bands A through G, but the upper band limit for band H is noticeably lower (£673,000 compared to £740,000). This results in a larger number of dwellings moving into bands H and I. Under the first method, 5,276 dwellings move from G to H and 1,796 from H to I. Under the 2005 revaluation method, we estimate these movements to be 6,450 and 2,893 respectively.

One consequence of adopting an alternative banding approach, but without altering the tax band multipliers, is that the various approaches are not revenue neutral but give rise to a different tax base (of course, this would then flow on to influence the calculation of Revenue Support Grant to local authorities). We emphasise that we do not advocate ignoring the tax multipliers. Indeed, amending them is an obvious policy option that could reduce the regressivity of the tax. But, to be clear, the analysis reported here deliberately assumes no changes to the multipliers in order to demonstrate the distributional impacts of the range of revaluation approaches. By calculating the total number of band D equivalents under each approach, we estimate the tax base as follows:

- Band limits that replicate the 2005 distribution: 1,409,200 (option 1)
- Band limits that replicate the 1993 distribution: 1,337,100 (option 2)

- Band limits set using the 2005 approach: 1,419,300 (option 3)
- Equal proportion of dwellings in 9 bands: 1,902,800 (option 4)
- Band limits set using cluster analysis: 1,601,400 (option 5)
- Current tax base for comparison: 1,402,750

The implications for individual local authorities are summarised in table 10 which sets out the percentage change in the total number of band D equivalents under the five options.

Table 10 Change in tax base arising from different banding approaches

Local authority	To replicate 2005 list (option 1)	To replicate 1993 list (option 2)	2005 revaluation method (option 3)	Equal proportion in each band (option 4)	Cluster analysis results (option 5)
Blaenau Gwent	-4.0%	-5.9%	-2.5%	2.1%	4.2%
Bridgend	0.7%	-4.0%	1.5%	34.9%	14.2%
Caerphilly	-2.0%	-5.9%	-0.5%	22.2%	9.0%
Cardiff	3.5%	-3.0%	3.7%	50.1%	19.6%
Carmarthenshire	-1.8%	-6.5%	-1.0%	30.5%	11.3%
Ceredigion	1.9%	-4.0%	2.2%	47.5%	17.3%
Conwy	3.1%	-2.8%	3.5%	46.0%	17.7%
Denbighshire	-0.1%	-5.4%	0.7%	37.8%	13.5%
Flintshire	-4.3%	-9.2%	-3.7%	32.7%	9.0%
Gwynedd	4.4%	-0.9%	5.2%	42.9%	18.9%
Isle of Anglesey	6.9%	0.9%	7.2%	51.3%	22.7%
Merthyr Tydfil	-6.3%	-8.5%	-4.7%	3.1%	3.1%
Monmouthshire	9.4%	1.5%	9.3%	60.6%	28.2%
Neath Port Talbot	-4.4%	-7.7%	-3.1%	14.4%	6.0%
Newport	0.1%	-4.9%	1.0%	35.4%	13.9%
Pembrokeshire	0.9%	-4.8%	1.3%	40.3%	15.4%
Powys	1.9%	-4.3%	2.2%	44.1%	17.2%
Rhondda Cynon Taf	-2.5%	-5.5%	-1.4%	16.6%	7.4%
Swansea	-1.3%	-6.1%	-0.3%	29.4%	11.9%
Torfaen	-2.7%	-6.8%	-1.3%	23.6%	9.0%
Vale of Glamorgan	6.4%	-0.8%	7.0%	50.5%	23.0%
Wrexham	-4.5%	-9.2%	-3.8%	27.9%	8.2%

The second column of table 10 (option 2) clearly shows the inflationary impact that the 2005 revaluation had on the council tax base. Returning to the 1993 distribution of dwellings by band would decrease the tax base in all but two local authorities (Isle of Anglesey and Monmouthshire) in Wales and by as much as 9.2% in Flintshire or

Wrexham. A pure revaluation (option 1) would result in a less dramatic change in the distribution of the council tax base, but some of the changes are still quite substantial. For example, Merthyr Tydfil, Wrexham, Flintshire, Neath Port Talbot and Blaenau Gwent would all lose more than 4% of their tax base. Monmouthshire, Isle of Anglesey and Vale of Glamorgan would realise the largest increase in their tax base. Please refer to the appendix for a demonstration of the proportion of dwellings predicted to move up, or down, by one council tax band, by local authority area.

The revalued upper band limits suggested by the cluster analysis (final column of table 10, option 5) reveal some interesting patterns. Although this approach leads to an increase in the tax base in all local authorities, in general, these band limits appear to have the greatest impact on local authorities that have seen the greatest increase in dwelling prices since 2003. The areas with the smallest increase in tax base include Blaenau Gwent, Caerphilly, Flintshire, Neath Port Talbot, Rhondda Cynon Taf, Torfaen and Wrexham (all areas gain less than 10%). The local authorities experiencing the largest gains include Cardiff, Isle of Anglesey, Monmouthshire and Vale of Glamorgan (all above 20% increase to their tax base).

Conclusions

This final report is intended to provide robust evidence to aid debate and exploration of the options for revaluing the council tax base in Wales. The report makes no policy recommendations and takes on no ideological position in relation to the appropriateness of the council tax, or options for replacement. However, the regressive nature of the council tax is well documented in a number of previous studies, and remains one of the principal criticisms of the tax. We note from our analysis that a simple revaluation would go some way towards addressing this issue if an alternative approach were taken to the design of band limits. In particular, when we used cluster analysis to reveal nine natural groupings of dwellings (based on value), our prediction was that the number of dwellings moving up bands better reflects the differential movement of prices within Wales since the last revaluation. Future research could usefully explore the likely impact of such a banding approach on households, taking account of household incomes. We note, for example, that other studies have used large scale survey datasets to estimate such impacts. However, such an analysis was beyond the scope of this study.

A key conclusion from this research is that statistical methods can be used to great effect, when combined with appropriate high quality data on dwelling attributes and transactions. The modelling exercises reported here resulted in the development of sophisticated, yet transparent and understandable, tools that can be used in the estimation or market value. The predictive accuracy of these tools is high, and arguably compares favourably with more time-intensive, traditional approaches. More manual, traditional, approaches will probably always be needed to reflect the valuation of atypical dwelling types, sizes or locations, and it is important to remember that hedonic methods work best when applied to frequently observed transaction types. Yet, we would argue that the approaches piloted and reported here indicate that statistical methods are an effective approach that could be used to allow more frequent and systematic monitoring of current property values in the future.

Our analysis of housing market trends in Wales since the last revaluation in 2005, and since the global financial crisis, shows that the timing of council tax revaluation could be very important indeed. We observe that the transaction prices of dwellings in lower council tax bands increased much more strongly in the run-up to the GFC than those in upper bands. The pattern of change since the GFC has been quite different, but there is a persistent trend of slightly higher price inflation in lower than upper bands. This, of course, suggests that the timing of revaluation and choice of base year are important, and should be selected to avoid short-term distortions caused by a boom or slump in housing market conditions. Arguably, our choice of 2019 as a hypothetical revaluation year might reflect such a period of stability. However, it is also worth pointing out that one of the undesirable attributes of the council tax is that infrequent revaluation leads, by definition, to more significant adjustments when revaluations do occur. The obvious counter to this would be for a system of more frequent and regular revaluations. There is danger that periods of instability, such as the boom conditions prior to the GFC, could result in greater short-term distortions to the distribution of dwellings by council tax band under a system of more frequent revaluation. On the other hand, it is also arguable that a system of more frequent revaluations could act as an automatic stabiliser (because rapidly increasing prices disproportionately affecting some localities or dwelling types would quickly be reflected in tax liability).

Finally, it should be noted that the full implications of the SARS-CoV2 (COVID-19) pandemic on housing prices, and the distribution of values, remained unknown at the time this study was carried out. However, it seems likely that a period of significant housing market adjustment could occur, and it would be prudent to consider the implications on the timing and choice of revaluation year should the Welsh Government decide to move towards a revaluation of the council tax base.

References

Adam, S., Hodge, L., Phillips, D and Xu, X. (2020) Revaluation and reform of council tax in Wales: impacts on different councils and household types, London: Institute for Fiscal Studies

Hills J and Sutherland H (1991) "The Proposed Council Tax", Fiscal Studies, 12, 4, 1-21.

Institute for Fiscal Studies (2011) The Mirrlees Review: Reforming the tax system for the 21st Century, London: Institute for Fiscal Studies.

Jones, C., Leishman, C. and Orr, A. (2006a) The revaluation of council tax bands: more than rearranging the deck chairs, Policy and Politics, Vol. 34, No. 2, 219-239.

Jones, C., Leishman, C. and Orr, A. (2006b) The potential impact of reforms to the essential parameters of the Council Tax, Fiscal Studies, Vol. 27, No. 2, 205-229.

Kenway P and Palmer G (1999) Council Tax: The Case For Reform, New Policy Institute, London.

Leishman, C. (2009) Spatial change and the structure of urban housing sub-markets, *Housing Studies*, Vol. 24, No. 5, 563-585.

Leishman, C., Bramley, G., Stephens, M., Watkins, D. and Young, G. (2014) After the Council Tax: impacts of property tax reform on people, places and house prices, York: The Joseph Rowntree Foundation.

Lyons, M. (2005) Lyons Inquiry into Local Government: Interim Report, London: The Stationery Office.

Muellbauer J and Cameron G (2000) Five Key Council Tax Reforms and Twelve Reasons to enact them, New Economy, June 2000, 7, 2, 88-91.

Muellbauer, J. (2005) Property taxation and the economy after the Barker Review, The Economic Journal, 115, pp. C99-C117.

Plimmer F, McCluskey W J and Connellan O (1999) "Reform of UK local government domestic property taxes", Property Management, 17, 4, 336-352.

Rajiv Prabhakar (2016) How did the Welsh government manage to reform council tax in 2005?, Public Money & Management, 36:6, 417-424, DOI: 10.1080/09540962.2016.1206750

Valuation Office Agency (2015), Council Tax Band Changes Due to 2005 Revaluation in Wales.

Appendix: proportion of dwellings predicted to move council tax bands

Local authority	(Option 1)		(Option 3)	
	% moving up 1 band	% moving down 1 band	% moving up 1 band	% moving down 1 band
Blaenau Gwent	6.7%	23.9%	11.3%	20.6%
Bridgend	21.1%	20.2%	24.3%	16.8%
Caerphilly	14.2%	22.6%	19.8%	17.6%
Cardiff	23.0%	15.3%	24.7%	14.2%
Carmarthenshire	16.6%	28.2%	19.0%	24.6%
Ceredigion	26.0%	18.8%	27.3%	17.6%
Conwy	25.4%	16.6%	26.7%	14.1%
Denbighshire	17.6%	21.3%	19.9%	17.6%
Flintshire	10.2%	32.6%	12.7%	28.5%
Gwynedd	27.5%	16.8%	30.2%	14.1%
Isle of Anglesey	36.9%	9.6%	39.5%	8.4%
Merthyr Tydfil	5.8%	27.1%	10.6%	24.0%
Monmouthshire	37.6%	8.9%	37.3%	9.1%
Neath Port Talbot	9.2%	32.9%	12.3%	27.1%
Newport	17.4%	22.7%	19.8%	18.4%
Pembrokeshire	22.0%	23.2%	23.4%	20.7%
Powys	23.4%	20.9%	24.6%	19.5%
Rhondda Cynon Taf	10.7%	21.3%	14.1%	18.3%
Swansea	12.9%	28.1%	15.6%	23.2%
Torfaen	12.2%	25.5%	18.0%	20.6%
Vale of Glamorgan	25.4%	15.3%	27.3%	13.1%
Wrexham	8.4%	37.5%	10.3%	33.4%