Article

Power Laws and Inequalities: The Case of British District House Price Dispersion

David Paul Gray ![ORCID]

Department of Accountancy Finance and Economics, University of Lincoln, Lincoln LN6 7TS, UK;
dgray@lincoln.ac.uk

Abstract: Descriptive statistics that are easy to generate and interpret are central to policy decision making. The GINI coefficient and the coefficient of variation are used widely when assessing inequality. In many areas of inequality, such as wealth and income holdings, the distribution is skewed. Here, simple power laws could provide useful 'descriptive' exponents. The Zipf-Pareto power law and Lavalette's law are used to reveal a steepening in the distribution of district house prices in Britain that began before the financial crash of 2008. The time profiles indicate the exponents closely mirror those of the GINI coefficient and the coefficient of variation. As such, they are useful tools in the quantification of inequalities.

Keywords: quintile ratio; house price in Britain; Lavalette's law; inequality; coefficient of variation

1. Introduction

In the economic growth literature, sigma-convergence is said to exist with a (persistent) narrowing in the spread of income levels. Standard measures of sigma-convergence are the standard deviation and the coefficient of variation (CoV). These are two of a battery of indicators of inequality that are used by agencies such as the OECD or the European Commission. A standard measure of inequality is the GINI coefficient. Both it and the CoV cover the full cross-section of the distribution.

Although International Territorial Level (ITL) (formerly NUTS) classifications are used across Europe, they are not of a standard size across countries, so comparing the degree of inequality in one country with another could produce misleading results. McCann (2020) was in dispute with a fact-finding website that challenged his finding that the UK was the most unequal country in Europe according to a number of measures. They used data at ITL2 and ITL3 levels when McCann used just ITL3. The measures he used were the GINI and CoV coefficients. This was supplemented with the ratios of the highest 10% [20%] GDP per capita regions to the lowest 10% [20%] GDP per capita regions. From here on, these are called the decile ratio (DR) and the quintile ratio (QR). They reveal 'slopes' within the distribution. These supplement distribution measures, unaffected by outliers.

Van Nieuwerburgh and Weill (2010) find that the U.S. regional house price distribution from the mid-1970s to the financial crisis became broader. Using a coefficient of variation (CoV) as the key indicator of dispersion, they argue that the driver is productivity differences. When investigating the spread in Hong Kong's residential prices, Leung et al. (2006) consider a broad range of macro variables including the growth rate of housing loans and the real GDP growth rate (credit and income). They use standard deviation and skew as measures of dispersion. They conclude that the two are associated with different aspects of the market. The latter is found to increase with interest rates and decrease with price level whereas the former is linked with housing loans, the stock market and the unemployment rate.

An observed regularity, expressed as Zipf's law, is a power law that posits a log-linear relationship between the rank-size of cities and their corresponding populations. Cristelli

et al. (2012) observe that Zipf's law has rapidly become an expression for measuring scale and size in many fields but has mixed empirical evidence for it. A Zipfian expression has also been used to assess sigma-convergence (Shao et al. 2011; Tang et al. 2016). If the coefficient tends to zero, there is a narrowing of the spread.

A Lavalettean power law has been found to be a better alternative for characterising tax income for Italian municipalities and regions than a Zipfian one (Cerqueti and Ausloos 2015a, 2015b), as well as for house-price–earnings ratios (Gray 2022). Contemporary econophysics asserts that the distributions of wealth and income holdings for all human societies have a composite statistical structure which is skewed and unimodal. It is log-normal in its bulk with power-law tails (Eliazar and Cohen 2013). Lavalette's expression can reflect distributions that are subject to skew and kurtosis but still are cumulatively *S*-shaped. As Zipfian exponents have been used to reveal convergence or changing inequalities, other power laws, such as a Lavalettean one, could be used. Cerqueti and Ausloos' preference for a Lavalettean over a Zipfian function could be explored with prices rather than income.

Using district house price data, which have a long upper tail in line with the assertion above, standard measures of inequalities are estimated. The time paths or profiles of ratios, the GINI coefficient and the CoV are compared over time. Using the GINI coefficient as a focus, its time paths are compared with those of the Lavalettean and a Zipfian exponents. Simple sub-divisions of the above are explored and results are compared with more aggregated data sets. In the process, the paper asks whether there has been an increase in house price dispersion. If there has been, is this a post-financial-crisis feature? Are there consistent inferences when the data are in a more aggregated form? These questions are addressed using a variety of measures of dispersion.

The paper is structured as follows. First, there is a discussion of the distribution of economic activity across space with an emphasis on inequalities and imbalanced growth models. Second, there is an explication of finance and relative house prices. Zipf-Pareto and Lavalette's expression are reviewed in the Section 4, along with the other measures of distribution mentioned above. The analysis shows that Lavalette's exponent provides a useful description of the distribution of data drawn from administrative units. Rather than the general rise in house price dispersion implied by Van Nieuwerburgh and Weill (2010), an increase in exponents is consistent with a steepening of the 'gradient' of measures over most, but not all, of the period, starting before the financial crisis. This can be seen in more aggregated data sets.

2. Economic Systems and Power Laws

Agglomeration economies underpin the returns from co-location in an urban environment. These Marshallian advantages enhance productivity, support higher wages, and attract in-migration to the metropolitan area. However, there are what are known as crowding costs that follow. The competition for land that comes with higher wages and in-migration leads to the bidding up of accommodation prices. In a spatial general equilibrium framework (Roback 1982), an individual agent, seeking to maximise their utility, will select a location based on these trade-offs. In such a framework, intra-urban price dispersion is expected to generate an inverse relationship between accommodation cost and distance from the Central Business District (CBD) (DiPasquale and Wheaton 1996). This is based on a common productivity (income) level. A more complex model discussed by McCann (2013) has four income groups (high × 2, medium and low). He posits that rich groups maximise utility in different ways. The commuting decision can be further complicated when featuring house size preferences, life cycle issues, amenities and conspicuous housing consumption. The extent of the city will be a function of commuting. Functional urban areas combine low-density rural areas surrounding a city whose labour market is highly integrated with that city. The incomes associated with the rural dweller are possibly generated in the local centre. Brown et al. (2015) show that English labour markets are characterised by a significant amount of long-distance commuting. They argue that the rural–urban interface can be understood as an extensive geographical space where

particularly higher-status migrants leave the high-density urban area to live in low density rural areas (Szumilo 2019). Indeed, Szumilo (2019) argues that the key drivers of changes in amenity values is differences in preferences and wealth of households. Given the space preferences of the higher status commuter, rural areas within a PUA, rich groups outbid lower income households for large spacious dwellings in rural areas. A productivity premium is observed by DEFRA (2020) who reports that that GVA per job is lower in what are 'Predominantly Rural' areas than in 'Predominantly Urban areas'. One would expect that houses beyond the PUA boundary would be less expensive, ceteris paribus.

Carlino and Saiz (2019) find that 'Beautiful Cities' disproportionally attract those with high human capital, driving local house prices up, particularly with supply-inelastic markets. Amenities in declining cities attract the more wealthy, displacing other income groups. The distance to the CBD appears unimportant. McCann's (2013) first income group drive up prices near amenities close to the city centre.

The relative size of urban centres is described as the size structure. Central place theory (Hsu 2012) posits that there are a variety of centres nested within a spatial system. Because of certain minimum thresholds, a central place evolves to provide services for smaller, surrounding centres. The theory predicts that there will be a decreasing number of increasingly large centres. This is captured empirically by Zipf's law, a relationship that can be modelled as a non-linear relationship between the rank-size of cities and their corresponding populations. Brakman et al. (2020, p. 304) summarise that, empirically, Zipf's law is not strictly adhered to at the continental level. It is more representative of populations at the national scale. That said, distributions of city sizes in the last 60 years are imperfectly characterised by Zipf. Cristelli et al. (2012) argue that many real systems do not show true power law behaviour because they are incomplete or inconsistent with the conditions under which one might expect power laws to emerge. A consequence of the coherence characterizing Zipfian sets is that, in general, Zipf's law does not hold for subsets or a union of Zipfian sets. Fontanelli et al. (2017) propose that the 'split-merge' process of administrative territory resizing can render initial power laws of cities into other kinds of probabilistic laws.

New Economic Geography models, applying agglomeration economies (Brakman et al. 2020), predict an imbalanced or dual regional economy. Using Verdoorn's law, which links output with productivity growth, a Thirlwallian regional growth model predicts constrained growth: the periphery has its growth rate constrained by that of the core (McCombie 1988). The periphery, often modelled as the agricultural sector in New Economic Geography, should have persistently lower productivity. UK regional productivity inequalities, which have been increasing since the early 2000s (Gal and Egeland 2018), are linked to the low productivity of UK regions and, in particular, some of its major cities. Persistently poor city performances have been attributed to disadvantageous industrial mixes, such as those linked with deindustrialisation, especially in the northern cities of England (Martin et al. 2018; Pike et al. 2016). The persistence has also been seen in the light of lock-in. Agglomeration economies trace out productivity development paths that are limited in trajectory. The economy's development path is dependent on previous investments, innovations, institutional arrangements, etc. that forged the current economic structures. These will shape future investment decisions (Martin et al. 2014, 2018). Thirlwall's regional growth model is trade-based. The weak growth rates of these northern cities are posited to be functions of their export structures and, in particular, low income-elastic output.

3. Finance and the Distribution of Local House Prices

DiPasquale and Wheaton (1996) assert that there are four factors that affect land prices. There is an inverse relationship between the cost of capital and asset prices. Over the period 1985 to 2018, which covers both the 1989 and the 2008 housing bubbles in the UK, there was a sustained decline in real interest rates (Miles and Monro 2019). Amaral et al. (2023) argue that the decrease in real interest rates has increased dispersion through inflating prices in major UK cities. Cyclical effects can be locked in to price. Bogin et al. (2017) find that

prices are higher than the 'pre-acceleration' level in larger centres, whereas in small cities, permanent price gains are not so clear. They conclude that the average growth acceleration is a signal of a persistent shift in a location's perceived economic fundamentals. This might explain the dislocation of capital cities or city regions from the rest. Himmelberg et al. (2005) argue that house prices in rapidly growing cities are more sensitive to changes in real interest rates.

The second factor proposed by DiPasquale and Wheaton is the expected future growth of current rent. Rental yields are expected to grow following a rise in population (Glaeser and Gyourko 2005) or income/productivity (Coulson et al. 2013; Van Nieuwerburgh and Weill 2010). In a well-functioning capital market, all assets should generate the same returns after adjusting for risk (Varian 2014, p. 244). Housing costs and house prices will be related, in part, to the lender's view of risk-in-lending. A spatial equilibrating process, based on risk-adjusted returns to mortgage lending across space (Sinai 2010), would imply a dispersion of prices across territories, in the face of anticipated changes. Areas with declining populations, low wage growth, etc., will be viewed as having poor credit scores, deterring lending there. Rae (2015) points to deprivation as influencing mortgage-lending appetite. Leung et al. (2006) provide a macro view of expected returns. The balance of payments and budget surplus/deficit herald future tax changes to correct imbalances.

The cost of housing for many buyers is such that they turn to a mortgage provider to address their financial constraints. Lenders can generate a financial accelerator (Aoki et al. 2004) through their lending practices. A positive shock leading to a rise in housing demand boosts prices. Those seeking to purchase and have a dwelling to sell benefit from an increase in their net worth. As they are perceived to be a lower risk, lenders could extend the amount of credit they are willing to offer, which boosts prices further, inflating house price–income ratios.

Using the now-standard Phillips and Sul process to reveal convergence clubs in UK house prices, Montagnoli and Nagayasu (2015) find the South East outer ring plus Northern Ireland, the Midlands and Northern regions are convergence clubs, suggesting common long-term price growth. Dublin and London may have decoupled from the rest of the British Isles (Gray 2018; Montagnoli and Nagayasu 2015; Richmond 2007).

Huang et al. (2015) seek to untangle the impact of amenities and credit on the Chinese experience of rising prices pre and post financial crisis. They find that the credit channel applies after the crisis, whereas amenities can explain relatively faster growth in house prices in the pre-crisis period.

4. Method

A rank-size function can be expressed as $P_R = P_1[R]^{-\alpha}$ (1) (Arshad et al. 2018) where P_R is the population of a city ranked as R and P_1 is that of the largest city. For Zipf's law to hold, the $-\alpha$ exponent, hereafter referred to as the Zipf-Pareto (Z-P) exponent, should be 1. Lavalette's ranking power law can be expressed as $P_R = P_1\left[\frac{N \times R}{N-R+1}\right]^{-q}$ (Gray 2022) (2), where N is the total number of cities. In both cases, P_1 is the calibrating value (CV). The formula describes a semi-logarithmic *S*-shape (Chlebus and Divgi 2007).

The cross-sectional models are estimated using simple OLS as $ln(P_R)_t = -\gamma_t - q_t ln\left(\frac{N \times R}{N-R+1}\right)_t$ and $ln(P_R)_t = -\delta_t - \alpha_t ln(R)_t$ for $R = 1 \ldots N, t \ldots T$. The intercept in both cases is the estimated CV in log form. The expected price for a given rank score can be generated and compared with the observed using a Kolmogorov–Smirnov test.

Standard sigma-convergence metrics include the CoV. This is defined as the cross-sectional standard deviation divided by the mean. Changes in descriptive statistics such as the expected mean, CoV and median are linear functions of the calibrating value for both (1) and (2). The GINI coefficient, defined as $\frac{1}{N}\left(N + 1 - 2\left(\frac{\sum_{R=1}^{N}(N-R+1)P_R}{\sum_{R=1}^{N}P_R}\right)\right)$, provides a standard measure of inequality. This takes a value between 0 and 1, where the closer to zero the coefficient is, the more even the distribution.

It is common to produce the CoV alongside within-distribution ratios that can assess a slope; the spread for central portion avoiding the tails. Two standard ones are the quintile ratio (QR) and the decile ratio (DR) used by, inter alios, McCann (2020) and Brandily et al. (2022). When $2q = \alpha$, the two functions generate the same lowest income with the same calibrating one. Given α, q, N and CV, one can generate expected values for QR and DR. A greater absolute q corresponds with more skew and fatter tails affecting the GINI coefficient, QR, DR and CoV.

5. Data

UK's Office for National Statistics (ONS) provides a repository for a number of house price data sets for England and Wales. Additionally, Statistics.gov.scot offers Scottish data at the district level. All the data are freely available. The median house price of 'all dwellings' in a year by Local Authority District is available for 339 districts in England and Wales plus 32 Scottish territories where the values are over £40,000. The Isles of Scilly data, which is intermittent, is excluded. Additionally, ONS provides travel-to-work areas data for England and Wales (171) plus a further two areas on the border between England and Scotland. At the ITL3 level, there are 145 territories. The data, which run from 2004 to 2019, are deflated using the retail price index back to 2004 levels. They cover the run up to, and the recovery from, the financial crisis of 2007-8. Figure 1 displays boxplots for all districts as well as the set bifurcated by district type classified as primary urban areas (PUAs) or cities and non-PUAs or rural areas. PUAs are the functional urban areas of the 62 major cities of Britain as defined by The Centre for Cities'[1]. These are covered by 146 districts, some of which will be less urban than others. In addition, there are 224 non-PUA districts.
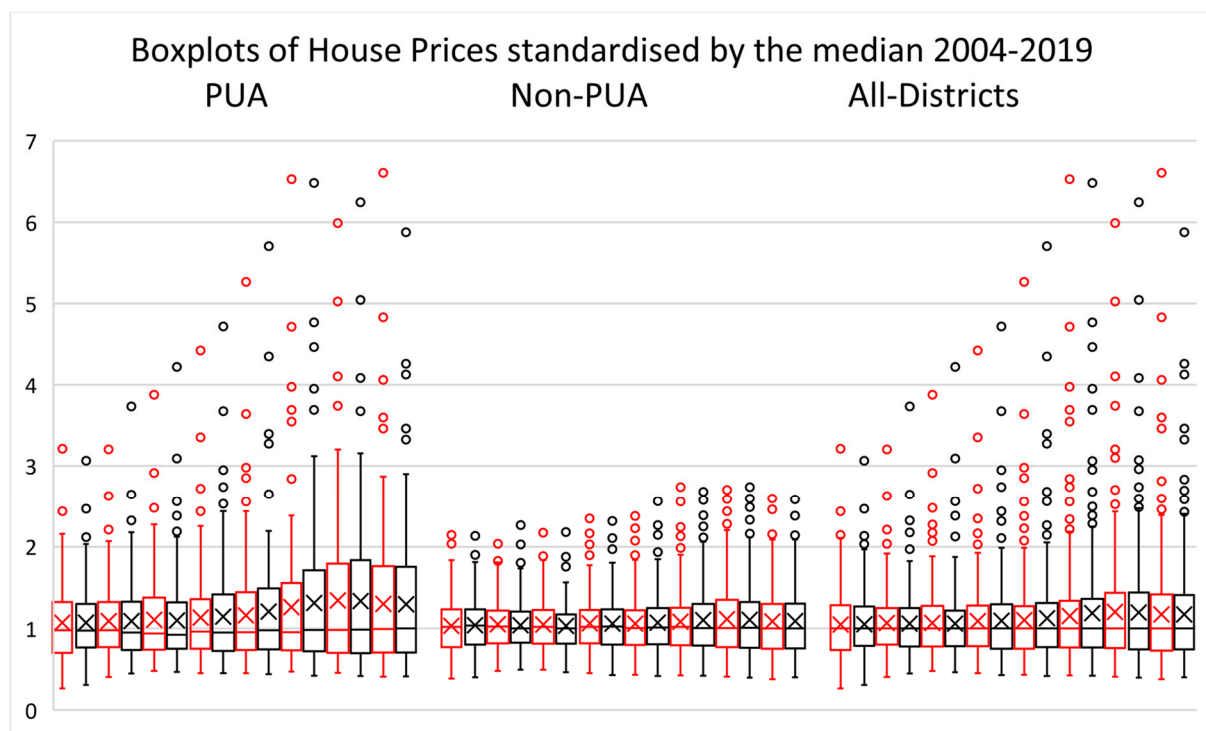

Figure 1. Boxplots of house price distributions by year and urban group.

The years' values are standardised using the median drawn from the complete set for the corresponding year. Figure 1 shows the outliers and longer whiskers at the top end for PUA districts in particular. The outliers do not change in line with the whiskers. The box expands, which would have a direct effect on the quartile ratio. Non-PUA districts have very stable boxes, whiskers and outliers with ranges much narrower than PUA districts.

Figure 2 displays time profiles of measures of inequalities. These measures of inequality are not necessarily of common value. To compare 'shapes', the profiles are indexed by the corresponding minimum value for non-PUAs in 2009, which provides the minimum value in all but the CoV.
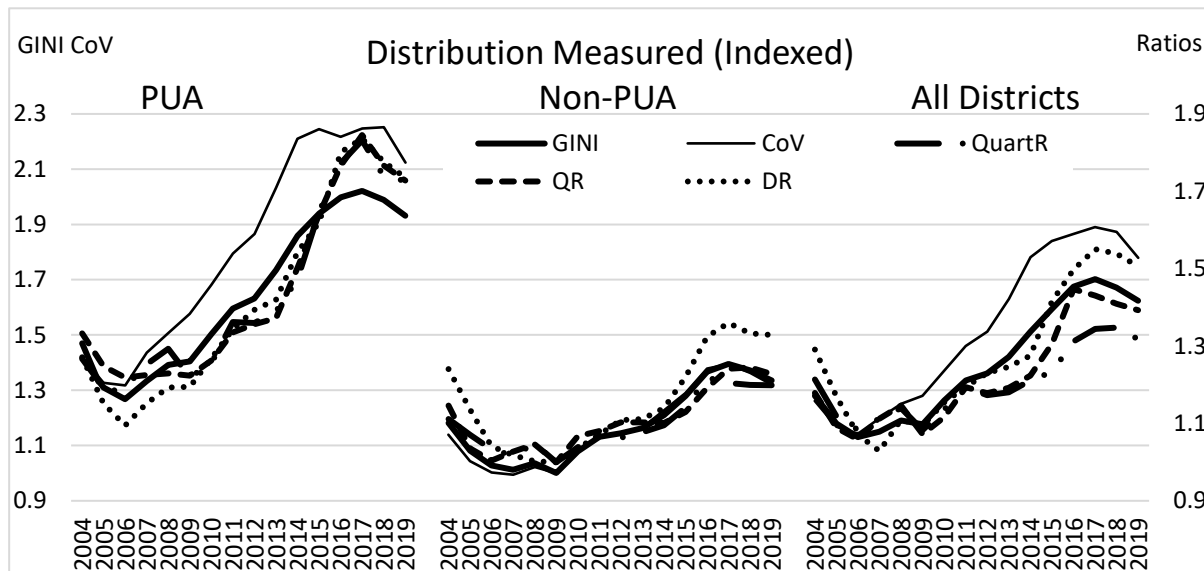

Figure 2. Distributional measures over time.

Reflecting the boxplots, the PUAs appear to have a wider range than non-PUA districts. All spread measures appear as *S*-shaped, declining in inequality to 2006/7 and then expanding to a peak around 2017. The CoV is more volatile than others, likely to reflect outliers seen in Figure 1, particularly with city districts. The PUA and non-PUA districts have a broader range of prices in 2017 compared with the pre-2008 peak in 2004, consistent with a steepening of the price distribution.

Generally, the Non-PUA series are the most similar to each other. Using the GINI coefficients as a reference, the CoV of the non-PUA districts almost duplicates the pattern in the index. With this benchmark, the CoV profiles for the PUAs and all districts are slightly higher than the GINI, suggesting greater dispersion over time. This could reflect the treatment of deviations, providing differing views of steepening.

The slope measures (righthand scale) show a similarity in both shape and proportional change across all three district collections. The DR is the most volatile, and the quantile ratio the least. The DR is the best aligned with the GINI values for non-PUA district dispersion. However, one might say that they have shallower profiles [all districts] or lower values [PUA] than the GINI benchmark.

The peak QR and quartile ratio values are distributed across 2016–2018, whilst the steepening might be said to start from 2009. This contrasts with the more obvious minimum in 2006 in the GINI and CoV evolution. The 2017 peak is not obvious in the CoV pattern for urban districts, which is flat from two or three years earlier, paralleling the outliers' patterns.

The estimates of the Lavalette's [and Z-P] exponents and CVs are reported in Table 1 for 2004 to 2019. The coefficients of determination (0.986 [0.775]) indicate Lavalette's expression is a better fit with the house price data. Kolmogorov–Smirnov test *p*-values support the view that a Lavalettean characterisation of house prices is not inappropriate. The CVs suggest real price levels are not dissimilar in 2012 to those in 2004. They trend upwards from 2009. This set of indicators suggests convergence or a decline in inequality to 2006 and then expanding to a peak around 2017.
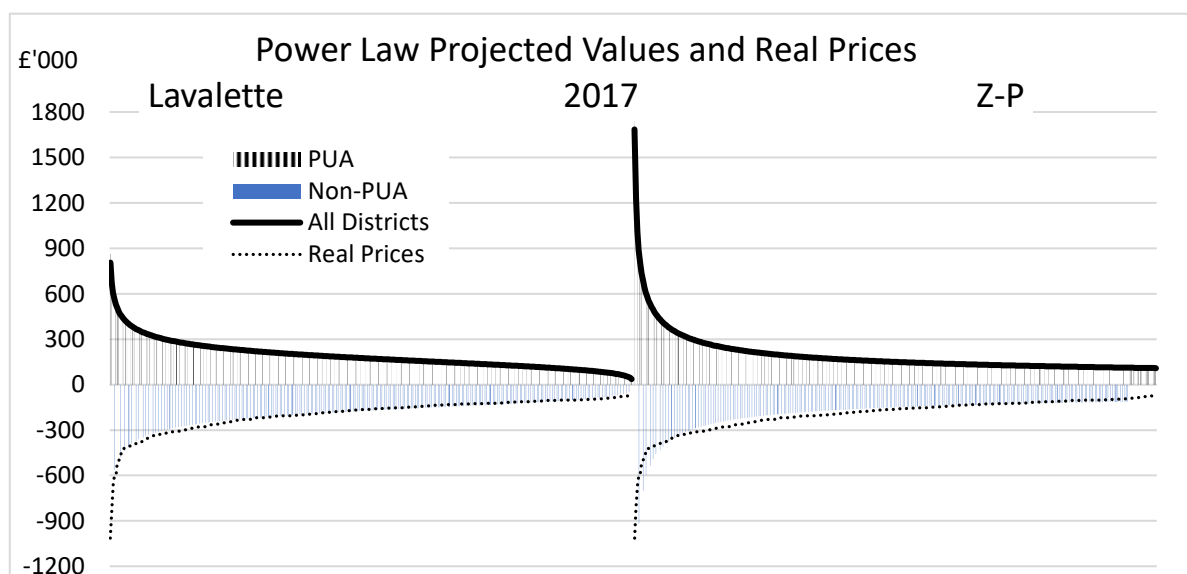
Table 1. Absolute exponent nationally and PUA/non-PUA.

	EWS q	R^2 q	CV q	K-S q p-Val.	EWS α	R^2 α	CV α	K-S α p-Val.	PUA q	PUA α	Non-PUA q	Non-PUA α
2004	0.215	0.986	487,030	0.472	0.349	0.775	759,075	<0.001	0.239	0.400	0.204	0.325
2005	0.196	0.991	458,939	0.774	0.321	0.798	700,390	<0.001	0.218	0.371	0.184	0.297
2006	0.183	0.991	437,908	0.774	0.309	0.849	680,811	<0.001	0.204	0.360	0.172	0.284
2007	0.183	0.984	460,314	0.53	0.316	0.880	739,517	<0.001	0.210	0.378	0.167	0.279
2008	0.187	0.977	462,056	0.24	0.326	0.893	763,492	<0.001	0.217	0.393	0.169	0.286
2009	0.185	0.979	412,394	0.53	0.325	0.909	686,198	<0.001	0.217	0.398	0.164	0.280
2010	0.198	0.981	472,642	0.278	0.348	0.909	813,374	<0.001	0.233	0.426	0.176	0.300
2011	0.208	0.977	492,685	0.418	0.367	0.913	877,254	<0.001	0.246	0.450	0.184	0.314
2012	0.211	0.975	497,050	0.32	0.373	0.915	897,356	<0.001	0.251	0.459	0.186	0.318
2013	0.218	0.971	520,783	0.278	0.390	0.926	974,644	<0.001	0.263	0.487	0.189	0.323
2014	0.229	0.963	580,549	0.206	0.413	0.940	1,146,962	<0.001	0.279	0.519	0.195	0.337
2015	0.242	0.964	664,024	0.206	0.435	0.933	1,349,860	<0.001	0.294	0.543	0.206	0.355
2016	0.255	0.964	751,260	0.088	0.456	0.922	1,568,640	<0.001	0.307	0.561	0.220	0.379
2017	0.261	0.967	806,241	0.105	0.463	0.912	1,685,219	<0.001	0.312	0.569	0.226	0.386
2018	0.258	0.971	802,638	0.149	0.455	0.905	1,640,559	<0.001	0.307	0.559	0.224	0.379
2019	0.252	0.973	775,617	0.24	0.443	0.900	1,547,266	<0.001	0.301	0.544	0.219	0.369

EWS: England Wales & Scotland; q drawn from the Lavalettean function (2), α drawn from a Z-P function (1). Kolmogorov–Smirnov test *p*-values.

Also reported are the exponents for the PUA and non-PUA districts. Importantly, the values for the non-PUA districts are lower and have a shallower time slopes but have the *S*-shape consistent with the actual measures seen in Figure 2, supporting a thesis of a rise in house price inequalities.

Figure 3 displays expected and actual prices for all districts for 2017. To illustrate the similarities, expected values for Lavalette and Z-P functions are shown as positive and real median district house prices are shown as negative values. The actual data are characterised by a long upper tail and an *S*-shape. The Lavalettean function (lefthand side) generates the *S*-shape, but the upper tail is shorter and the lower is more pronounced than the actual price data. The Z-P function (righthand side) predicts a greater CV with more change than the Lavalettean alternative. The CV in 2017 was 2.46× that in 2009, as opposed to doubling, as in the Lavalettean case. The Z-P power law models an L-shaped price hierarchy. The distinctive lower tail over-predicts the minimum price, whilst the CV is larger than the actual peak values.

Figure 3. Displays the exponents for all districts.

It is argued that Zipf 's Law is the ultimate signature of an integrated system (Cristelli et al. 2012). Lack of coherence or economic integration explains the failure of the data to reflect the law. Economic systems at the 'world' level represent the 'wrong scale'. Cities will not be fully integrated within a spatial system, explaining why city size will not follow a Zipfian process. Analysis at the national level is the right scale. Even then, London as a primary node that is oversized can be explained through the observation that its service role is broader than just the UK's. In house prices, it and Dublin may have decoupled from the rest of the British Isles, because of a broader financial services role.

Removing subsets of a Zipfian distribution and expecting the remaining data to still conform to the law is flawed. Indeed, common exponents should not be expected across the subdivision. However, the process here explores what impact a subdivision might have for coefficients and exponents. The Lavalette coefficients for the two subgroups (PUA values CV £864,871, $-q$ 0.312, R^2 0.936; non-PUA values CV £562,624, $-q$ 0.226, R^2 0.986) for 2017 reveal the former has a higher CV and a steeper distribution. Fusing these two together generates a third series of 370 values. Also in Figure 3 are stripes highlighting expected prices for primary urban areas, or cities (positive) and non-PUAs, or rural areas (negative). The layout entails how the third series weaves the two sets of data together by rank order so the spacing between stripes varies. As they have greater high values, reflecting the CV, the PUA stripes are denser in the lower ranks. The corresponding all-district values are reported in Table 1. Combining the corresponding Z-P values also produces something that reflects the expected values of the full series.

As both subset distribution tails should be appropriate (one short, one long and *S*-shaped), coherence (Cristelli et al. 2012) is more of an issue for the Lavalettean function. Even with a narrower spread, the maximum value of the non-PUA district distribution is lower than expected for both power laws. This reflects the scaling problem reported in Cristelli et al. (2012). This has implications for how one might interpret the values. The similar shape could reflect the coherence requirement (Cristelli et al. 2012). It should be the case that non-PUA house prices are linked to PUAs' through common risk management and mortgage lending policies, broad economic overspills, or arbitrage, so the range in the latter is influenced by the former.

6. Discussion

6.1. Median and Growth

Van Nieuwerburgh and Weill's steepening can be analysed using both the calibrating value and the exponent. Figure 4 displays the time paths of the Z-P and Lavalette exponents for all districts and for the urban–rural subdivisions. Both power functions indicate the spread in 2017 was the broadest and the narrowest in 2006 for three groupings, pointing to a steepening house price distribution in accord with Van Nieuwerburgh and Weill. The PUA exponent changed more than its non-PUA counterpart. Also, the Lavalettean calibrating value for PUA districts grew by 82%, faster than the 50% for non-PUA districts. The greater range of prices can be attributed to both intra- and inter-group factors. A core–periphery model, such as a Thirlwallian constrained growth model, posits that the core would have a greater growth rate and is more likely to have outliers at the top end, implying a wider range of growth rates than the periphery.

As well as the Z-P and Lavalette's exponents for the 16 years, Figure 4 also features the median district prices from the raw data and the expected medians generated from those exponents. They are in real terms. The actual growth of median district prices of 19% (PUA, 22%; non-PUA, 17%) is lower than the corresponding Lavalette or Z-P projections. The pattern, though, is consistent across all three collections. There is a rise in the median until 2007, followed by a collapse around 2009. It then undulates until 2012/2013 and then rises. 'Peak median' occurs in 2018. In the first period until 2009, generally, the exponents and median price moved in the opposite direction. After that, they are generally directly aligned. The scheme diverged.
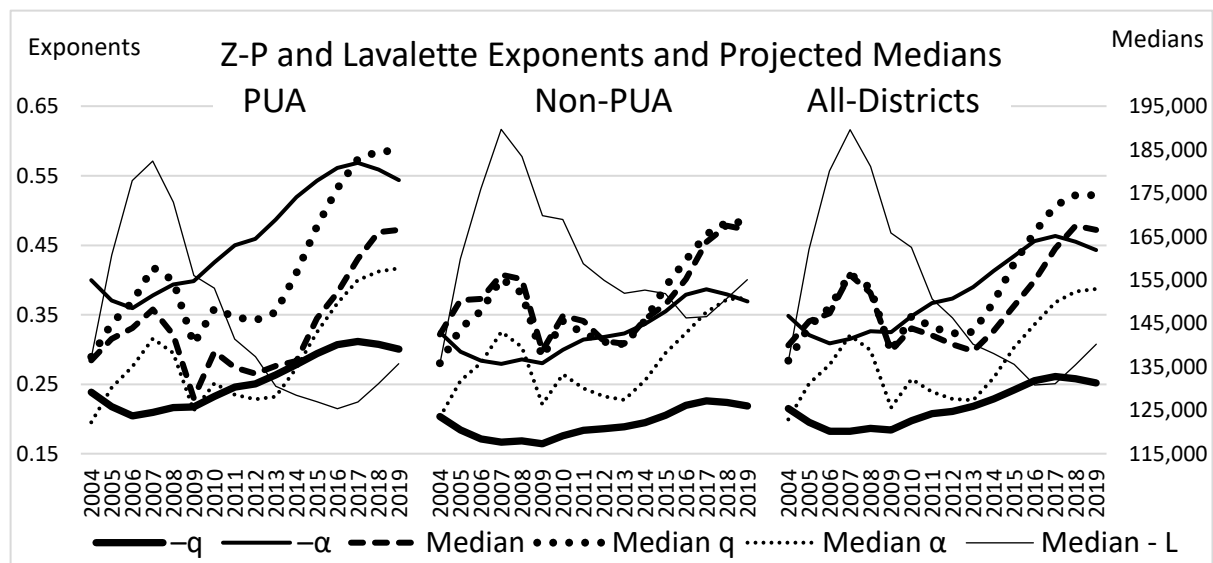
Figure 4. Power law exponents and projected medians.

For districts as a whole and for non-PUA areas, the growth rate at the lower quintile level was faster than at the corresponding median, implying convergence/a narrowing of inequalities. So, of the three, the median district has the most sluggish growth compared with the lower and upper quintile district. It is concluded that, given other measures, price growth at the median is slower than expected. This inference of convergence is at odds with all the measures applied in the analysis here.

Also in Figure 4 is a profile of the projected medians if they grew at the same rate as the implied lowest-priced district (Median-L). If the spread grows faster than the CV, the minimum value could fall. In the case of the city district distribution, the Lavalettean expression implies that districts at the bottom end were cheaper in real terms in 2019 than in 2004, despite the 82% increase in the most expensive district. The greater range has something to do with city performance at the bottom end. Core–periphery models would not predict this.

The much broader range in PUA districts reflects comments about productivity and income growth elsewhere. A persistently poor performance accords with Martin et al. (2018) and Pike et al. (2016). Indeed, two-thirds of large UK cities are less prosperous than the UK average (McCann 2020). The productivity growth of many of these has almost completely stalled since the 2008 crisis.

6.2. Comment on Aggregation and Time Profiles

The theoretical implications of aggregation are ambiguous. Hanushek et al. (1996) found that omitted variable bias increases with the aggregation of education data from school to state level. McCann (2020) reviewed the thesis of *The Economist* magazine that the UK is more unequal than any other industrialized country. This he did using ITL3-level productivity data. The question was whether *The Economist*'s results would change using different data aggregations. Merging areal units should pull the resulting unit value closer to the average, narrowing the range (maximum: minimum).

Aggregation effects are explored on three sets of house price data sets for England and Wales to which the two power functions can be applied. Travel-to-work areas and ITL3-level data are viewed as aggregated relative to the 338-district data. The criteria used for defining TTWAs suits a delineation of housing markets in the style of DiPasquale and Wheaton (1996), where a high proportion of the resident workforce works in that area. It is similar in concept to a FUA. ITLs and districts reflect administrative rather than economic units. Broadly, at the ITL3 level, there is an emphasis on merging areal units into 'urban' and 'hinterland', such as Derby plus two other areas of East Derbyshire and South and

West Derbyshire. TTWAs would combine elements of both urban and commutable rural areas, such as the TTWA of Derby.

The time profiles in Figure 5 are displayed in pairs. The GINI (first three profiles) and CoV (second three profiles) use the scale on the lefthand side. Both provide a similar picture of changing inequalities in the raw house price data. Interestingly, the ITL data have a slightly broader spread than the more disaggregated district-level data. However, the TTWA data have a notably lower range. Thus, aggregation can but does not necessarily lead to lower dispersion.
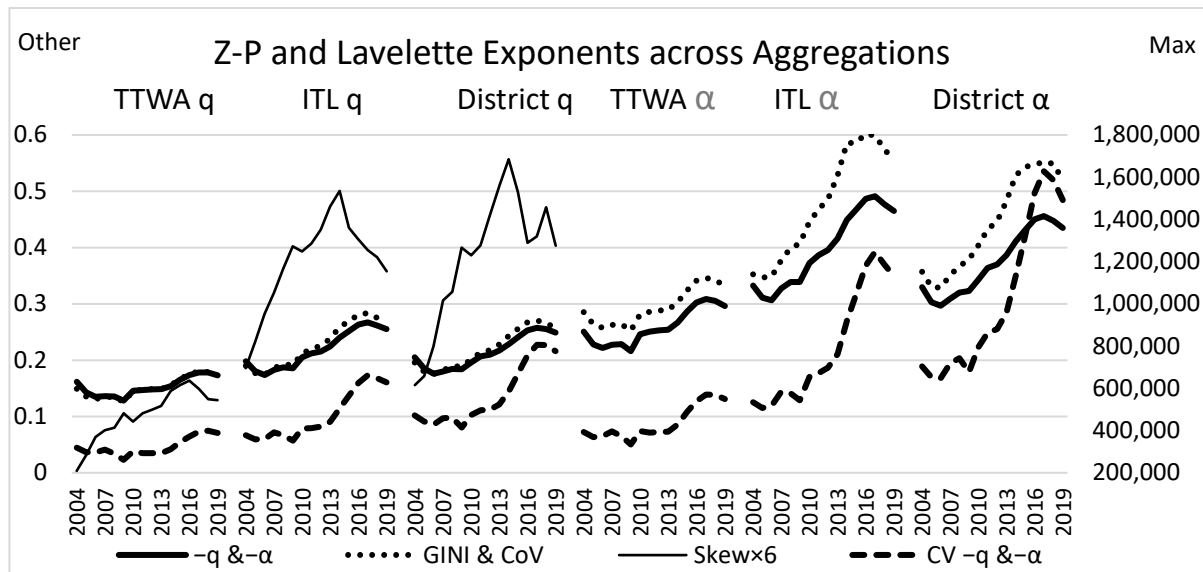

Figure 5. Power law exponents across aggregations.

The Lavalette (first three profiles) and Z-P exponents (second three profiles) can be compared with the GINI and CoV values. The GINI coefficient is shadowed by the Lavalettean exponent, whereas the Pareto values are larger. All show the *S*-shape seen above. Table 2 reports the CoV and slope ratios for the peak and trough years. There are some interesting consistencies. The district CoV value of 0.546 in 2017 is similar to the corresponding one for ITL3 (0.598). The slope ratios are also alike, implying they follow similar dispersions. This reinforces the point that aggregation can but does not necessarily lead to a lower dispersion. District slope observations in 2006 are slightly above those of TTWA data in 2017, yet the CoV is smaller. The CoV would be affected by outliers, which are large in 2017 at the district level.

Also shown in Figure 5 are the CVs in real terms. The district maximum values are greater than those of the ITL3 areas and the TTWA territories. Indeed, the expected max: min ratio (see Appendix A for the expression) is, for the former, 20.1, and for the latter, 15.7. Even though the range is widest at the district level, dispersion is not.

Using bootstrap methods, the confidence intervals for the exponents are estimated. As shown in Table 2, the 95% confidence interval for TTWA using the Lavalette estimator has the [absolute] upper bound of 0.139 in 2006. The corresponding lower bound value in 2017 was 1.44. From this, the two estimated exponents are deemed to be significantly different, supporting the claim that the price distribution has steepened. The same inference can be made for the other two data sets. This is also reflected in the results using the Z-P expression. The K-S statistic again suggests that the house price distributions generally follow a Lavalette function. However, that is not the case with district data [$p = 0.000$] in the peak year of 2017.

The upper- and lower-bound exponents estimated using bootstraps can be used to project what the slope ratios could be if the data followed power laws. In 2006, the DR for the TTWA data was 1.792. This falls within the 1.771 to 1.842 range, consistent with

Lavalette's law. There are other cases, exclusively with a Lavalette-derived decile ratio. Otherwise, the projected ratios are smaller than the observed ones. Leung et al. (2006) use standard deviation and skew in their analysis of house price dispersion. Their findings are that the two provide differing implications. The substitute for the standard deviation, the CoV, captures the change in spread in much the same way as the GINI coefficient and the slope ratios. However, skew time profiles, which are shown in Figure 5 multiplied by six to show variation using the lefthand-side scaling, do not follow the *S*-shape that all the other indicators show. There is no dip towards 2006. The peak is in 2014, not 2017. The aggregation does alter the time profiles, but the general shape remains the same. From this, it is not evident that skew alone provides insights into any steepening of the price distribution that is consistent with other measures. That said, Table 2 reports that the ITL3 and district data have high skew and kurtosis, which are significant. Skew and particularly kurtosis could explain why the quartile and quintile ratios are larger than those ratios associated with the power laws.

Table 2. Aggregation data and dispersion indications.

	K-S *p*-Val.		Coefficient	Quartile	QR	DR	Skew	Kurt
TTWA 2006		CoV	0.257	1.423	1.527	1.792	0.38	−0.16
Lavalette	0.876	Min	0.13	1.331	1.434	1.771		
		Max	0.139	1.357	1.470	1.842		
Z-P	0.002	Min	0.198	1.243	1.316	1.545		
		Max	0.26	1.331	1.434	1.771		
TTWA 2017		CoV	0.346	1.509	1.718	2.139	0.89	0.48
Lavalette	0.881	Min	0.166	1.440	1.584	2.074		
		Max	0.181	1.488	1.652	2.215		
Z-P	0.038	Min	0.285	1.368	1.485	1.871		
		Max	0.338	1.450	1.598	2.102		
ITL3 2006		CoV	0.348	1.627	1.832	2.263	1.69	4.7
Lavalette	0.623	Min	0.171	1.456	1.607	2.120		
		Max	0.186	1.505	1.675	2.265		
Z-P	0.003	Min	0.28	1.360	1.474	1.850		
		Max	0.352	1.472	1.629	2.167		
ITL3 2017		CoV	0.598	1.957	2.373	3.332	2.37	7.93
Lavalette	0.214	Min	0.246	1.717	1.978	2.948		
		Max	0.288	1.883	2.222	3.545		
Z-P	0.052	Min	0.463	1.663	1.900	2.766		
		Max	0.531	1.792	2.088	3.211		
LAD 2006		CoV	0.328	1.534	1.758	2.129	1.35	3.97
Lavalette	0.66	Min	0.173	1.462	1.616	2.139		
		Max	0.178	1.479	1.638	2.186		
Z-P	0.000	Min	0.278	1.357	1.470	1.842		
		Max	0.324	1.428	1.567	2.038		
LAD 2017		CoV	0.546	1.969	2.357	3.308	2.52	11.42
Lavalette	0.000	Min	0.249	1.728	1.994	2.987		
		Max	0.267	1.798	2.097	3.233		
Z-P	<0.001	Min	0.43	1.604	1.815	2.572		
		Max	0.491	1.715	1.975	2.941		

The thesis that London's housing market is dislocating from the rest of Britain's would lead to a steepening of the price distribution at the district and ITL3 levels of aggregation, where each of the 33 boroughs of London are included as separate housing areas. This dislocation is consistent with a combination of spatial sorting, leading to a concentration of those with high human capital, who are perceived as being high earners in the long run and so deemed lower-risk mortgage clients in the long run. The upshot is that London's and associated housing markets become marked by very high house prices. Exaggerating

this is the safe haven status of London, which attracts international investors in property, not to rent but just hold, as one might hold gold, as a financial asset (Fernandez et al. 2016).

7. Conclusions

In this paper, the time profiles of cross-sectional GINI and CoVs coefficients, which are used routinely as measures of inequality and sigma-convergence, are compared with exponents from two simple expressions. The Zipf-Pareto power law has been applied in a variety of ways (Cristelli et al. 2012) including to that of sigma-convergence (Shao et al. 2011; Tang et al. 2016). A power law due to Lavalette presents an alternative characterisation that fits a class of data with a long tail found in other areas of economic inequalities to be useful (Cerqueti and Ausloos 2015a, 2015b). All four series indicate that over the 2004–2019 interval there are two periods of narrowing of real district house prices on either side of a long period of broadening. The distribution of price in 2017 is broader than in 2006, as assessed using the four measures. As such, the time-path of the four indicate sigma-divergence between 2006 and 2017 and a Van Nieuwerburgh and Weill steepening of price. An explanation for this combines risk appetite and productivity arguments plus spatial sorting. London and the South East areas have very high prices and very high house price–earning ratios. The former could be explained by wages and so productivity alone. However, house price–earnings ratios are related to local risk assessments. Lenders must view borrowers in London as presenting no more risk, even when the loans are much higher when adjusting for the borrower's current income. The special nature of the borrower could be explained by London's attraction of those with the highest human capital. London is distinct from the rest of Britain. This suggestion would apply to other capitals/financial centres around the globe.

The power laws predict smaller slope ratio values than the observed ones, particularly the quintile ratio. It is likely that the central part of the price distribution is more concentrated than the exponents imply, possibly because of 'convergence' at the lower end of the distribution. Because growth at the median is modest, there is catchup from the bottom end, leading to greater kurtosis.

The subdivision into urban and rural show that the calibrating values and exponents capture a steepening of the 'gradient' of the district price distribution. The Lavalettean expression generates a more appropriate *S*-shaped distribution than the Z-P. Goodness of fit also favours Lavalette.

One of the issues raised by McCann (2020) is that one can make conflicting claims about inequality using differing levels of spatially aggregated economic data. Housing economics predicts that the high-human-capital worker would choose to live in areas that may be designated as separate from where the work is. TTWAs and primary urban areas would combine the two. When district and ITL data distinguish urban from rural administrative areas, this may separate the two. As such, disaggregation below the TTWA level could/should lead to widening inequality.

The district and ITL data provide a similar picture of absolute and changing inequalities of the raw house price data. Interestingly, the ITL data have a slightly broader dispersion than the more disaggregated district level. Thus, aggregation can, but does not necessarily, lead to lower dispersion. Given the nature of the aggregation, the notably lower dispersion found with TTWA data is to be expected. However, the time profile does not undermine the notion that there is a steepening of the price distribution in England and Wales.

Looking forward, the notion of levelling up, although a nice political slogan, could be better nuanced. As an exponent provides the expected distribution and slope ratios, an acceptable level of inequality could be established through setting a dispersion level, with the corresponding slope ratios highlighting where intervention is required. For instance, where house prices in some territory move beyond an acceptably high level, possibly threatening to dislocate from the rest, financial regulators could impose more stringent limits on lending metrics.

Appendix A

Table A1. Useful ratios and approximations.

		Zipf-Pareto	Lavalette
Expression		$P_R = P_1 R^{-\alpha}$	$P_R = P_1 \left(\frac{NR}{N-R+1} \right)^{-q}$
10th City by size compared with the 15th City	$\alpha = 1$ $q = 1$ and $N = 146$ $q = 1$ and $N = 224$	1.5	1.557 1.536
10th City by size compared with the 15th City	$\alpha = 0.25$ $q = 0.25$ and $N = 146$ $q = 0.25$ and $N = 224$	1.107	1.117 1.113
maximum ÷ minimum		N^{α}	N^{2q}
maximum ÷ median		$([N+1]\div 2)^{\alpha}$	N^{q}
Median		$P_1([N+1]\div 2)^{-\alpha}$	$P_1 N^{-q}$
Slope ratio	Quartile ratio would be assessed as $\omega = 75\%$; QR 80%; and DR at 90%	$\left(\frac{100-\omega}{\omega} \right)^{-\alpha}$	$\left(\frac{100-\omega}{\omega} \right)^{-2q}$
	$\alpha = 0.5$ $q = 0.25$	Quartile ratio = 1.732 QR= 2, DR = 3	Quartile ratio = 1.732 QR= 2, DR = 3
	90:50 ratio	$\left(\frac{10}{50} \right)^{-\alpha}$	$\left(\frac{10}{90} \right)^{-q}$
Coefficient of Variation. If the data follows a t-distribution	CoV of 0.398 is associated with a DR of 3.1; a QR of 2.01; and a Quartile ratio of 1.73.		

Note

[1] https://www.centreforcities.org/wp-content/uploads/2016/01/2016-PUA-Table.pdf (accessed on 3 March 2023).

References

Amaral, Francisco, Martin Dohmen, Sebastian Kohl, and Moritz Schularick. 2023. Interest Rates and the Spatial Polarization of Housing Markets. CEPR Discussion Paper No. 17780. American Economic Review: Insights, Forthcoming. Available online: https://www.aeaweb.org/articles?id=10.1257/aeri.20220367 (accessed on 7 July 2023).

Aoki, Kosuke, James Proudman, and Gertjan Vlieghe. 2004. House prices, consumption and monetary policy: A financial accelerator approach. *Journal of Financial Intermediation* 13: 414–35. [CrossRef]

Arshad, Sidra, Shougeng Hu, and Badar Nadeem Ashraf. 2018. Zipf's law and city size distribution: A survey of the literature and future research agenda. *Physica A: Statistical Mechanics and Its Applications* 492: 75–92. [CrossRef]

Bogin, Alexander, William Doerner, and William Larson. 2017. Local House Price Paths: Accelerations, Declines, and Recoveries. *The Journal of Real Estate Finance and Economics* 58: 201–22. [CrossRef]

Brakman, Steven, Harry Garretsen, and Charles van Marrewijk. 2020. *An Introduction to Geographical and Urban Economics: A Spiky World*. Cambridge: Cambridge University Press.

Brandily, Paul, Mimosa Distefano, Hélène Donnat, Immanuel Feld, Henry Overman, and Krishan Shah. 2022. *Bridging the Gap: What Would It Take to Narrow the UK's Productivity Disparities?* London: Resolution Foundation, June 30.

Brown, David, Tony Champion, Mike Coombes, and Colin Wymer. 2015. The Migration-commuting nexus in rural England. A longitudinal analysis. *Journal of Rural Studies* 41: 118–28. [CrossRef]

Carlino, Gerald, and Albert Saiz. 2019. Beautiful City: Leisure Amenities and Urban Growth. *Journal of Regional Science* 59: 369–408. [CrossRef]

Cerqueti, Roy, and Marcel Ausloos. 2015a. Evidence of economic regularities and disparities of Italian regions from aggregated tax income size data. *Physica A: Statistical Mechanics and Its Applications* 421: 187–207. [CrossRef]

Cerqueti, Roy, and Marcel Ausloos. 2015b. Cross ranking of cities and regions: Population versus income. *Journal of Statistical Mechanics: Theory and Experiment* 2015: P07002. [CrossRef]

Chlebus, Edward, and Gautam Divgi. 2007. A Novel Probability Distribution for Modeling Internet Traffic and its Parameter Estimation. Paper presented at the IEEE GLOBECOM 2007—IEEE Global Telecommunications Conference, Washington, DC, USA, November 26–30; pp. 4670–4675. Available online: https://ieeexplore.ieee.org/document/4411796 (accessed on 30 March 2020).

Coulson, N. Edward, Crocker H. Liu, and Sriram V. Villupuram. 2013. Urban economic base as a catalyst for movements in real estate prices. *Regional Science and Urban Economics* 43: 1023–40. [CrossRef]

Cristelli, Matthieu, Michael Batty, and Luciano Pietronero. 2012. There is more than a power law in Zipf. *Scientific Reports* 2: 812. [CrossRef]

DEFRA. 2020. *Statistical Digest of Rural England*; London: Department for Environment, Food & Rural Affairs. Available online: https://www.gov.uk/government/statistics/statistical-digest-of-rural-england (accessed on 13 October 2020).

DiPasquale, Denise, and William Wheaton. 1996. *Urban Economics and Real Estate Markets*. Englewood Cliffs: Prentice Hall.

Eliazar, Iddo, and Morrel H. Cohen. 2013. On the physical interpretation of statistical data from black-box systems. *Physica A: Statistical Mechanics and Its Applications* 392: 2924–39. [CrossRef]

Fernandez, Rodrigo, Annelore Hofman, and Manuel B. Aalbers. 2016. London and New York as a Safe Deposit Box for the Transnational Wealth Elite. *Environment and Planning A* 48: 2443–61. [CrossRef]

Fontanelli, Oscar, Pedro Miramontes, Germinal Cocho, and Wentian Li. 2017. Population Patterns in World's Administrative Units. *Royal Society Open Science* 4: 170281. [CrossRef]

Gal, Peter, and Jagoda Egeland. 2018. Reducing Regional Disparities in Productivity in the United Kingdom. Working Paper 1456, OECD Economics Department, Paris, France. [CrossRef]

Glaeser, Edward L., and Joseph Gyourko. 2005. Urban Decline and Durable Housing. *Journal of Political Economy* 113: 345–75. [CrossRef]

Gray, David. 2018. Convergence and divergence in British housing space. *Regional Studies* 52: 901–10. [CrossRef]

Gray, David. 2022. Do House Price-Earnings Ratios in England and Wales follow a Power Law? An Application of Lavalette's Law to District Data. *Environment and Planning B: Urban Analytics and City Science* 49: 1184–96. [CrossRef]

Hanushek, Eric A., Steven G. Rivkin, and Lori L. Taylor. 1996. Aggregation and the Estimated Effects of School Resources. *The Review of Economics and Statistics* 78: 611–27. [CrossRef]

Himmelberg, Charles, Christopher Mayer, and Todd Sinai. 2005. Assessing High House Prices: Bubbles, Fundamentals and Misperceptions. *Journal of Economic Perspectives* 19: 67–92. [CrossRef]

Hsu, Wen-Tai. 2012. Central Place Theory and City Size Distribution. *The Economic Journal* 122: 903–32. [CrossRef]

Huang, Daisy J., Charles K. Leung, and Baozhi Qu. 2015. Do Bank Loans and Local Amenities Explain Chinese Urban House Prices? *China Economic Review* 34: 19–38. [CrossRef]

Leung, Charles Ka Yui Leong, Youngman Chun Fai Wong, and Siu Kei. 2006. Housing Price Dispersion: An Empirical Investigation. *The Journal of Real Estate Finance and Economics* 32: 357–85. [CrossRef]

Martin, Ron, Ben Gardiner, and Peter Tyler. 2014. *The Evolving Economic Performance of UK Cities: City Growth Patterns, 1981–2011*; Working Paper, Foresight Programme on the Future of Cities; London: UK Government Office for Science, Department of Business, Innovation and Skills.

Martin, Ron, Peter Sunley, Ben Gardiner, Emil Evenhuis, and Peter Tyler. 2018. The city dimension of the productivity growth puzzle: The relative role of structural change and within-sector slowdown. *Journal of Economic Geography* 18: 539–70. [CrossRef]

McCann, Philip. 2013. *Modern Urban and Regional Economics*, 2nd ed. Oxford: Oxford University Press.

McCann, Philip. 2020. Perceptions of regional inequality and the geography of discontent: Insights from the UK. *Regional Studies* 54: 256–67. [CrossRef]

McCombie, John. 1988. A synoptic view of regional growth and unemployment: II The post-Keynesian theory. *Urban Studies* 25: 399–417. [CrossRef]

Miles, David, and Victoria Monro. 2019. *UK House Prices and Three Decades of Decline in the Risk Free Real Interest Rate*. Staff Working Paper No. 837. London: Bank of England.

Montagnoli, Alberto, and Jun Nagayasu. 2015. UK house price convergence clubs and spillovers. *Journal of Housing Economics* 30: 50–58. [CrossRef]

Pike, Andy Pike, Danny MacKinnon, Mike Coombes, Tony Champion, David Bradley, Andrew Cumbers, Liz Robson, and Colin Wymer. 2016. *Uneven Growth: Tackling City Decline*. York: Joseph Rowntree Foundation. Available online: https://www.jrf.org.uk/report/uneven-growth-tackling-city-decline (accessed on 30 March 2020).

Rae, Alasdair. 2015. The illusion of transparency: The geography of mortgage lending in Great Britain. *Journal of European Real Estate Research* 8: 172–95. [CrossRef]

Richmond, Peter. 2007. A roof over your head; house price peaks in the UK and Ireland. *Physica A: Statistical Mechanics and Its Applications* 375: 281–87. [CrossRef]

Roback, Jennifer. 1982. Wages, rents, and the quality of life. *Journal of Political Economy* 90: 1257–78. Available online: https://www.jstor.org/stable/1830947 (accessed on 30 March 2020). [CrossRef]

Shao, Jia, Plamen Ch. Ivanov, Branko Urošević, H. Eugene Stanley, and Boris Podobnik. 2011. Zipf rank approach and cross-country convergence of incomes. *Europhysics Letters* 94: 48001.

Sinai, Todd. 2010. Feedback Between Real Estate and Urban Economics. *Journal of Regional Science* 50: 423–48. [CrossRef]

Szumilo, Nikodem. 2019. The spatial consequences of the housing affordability crisis in England. *Environment & Planning A* 51: 1264–86. [CrossRef]

Tang, Pan, Ying Zhang, Belal E. Baaquie, and Boris Podobnik. 2016. Classical convergence versus Zipf rank approach: Evidence from China's local-level data. *Physica A: Statistical Mechanics and Its Applications* 443: 246–53. [CrossRef]

Van Nieuwerburgh, Stijn, and Pierre-Olivier Weill. 2010. Why Has House Price Dispersion Gone Up? *The Review of Economic Studies* 77: 1567–606. [CrossRef]

Varian, Hal R. 2014. *Intermediate Microeconomics: A Modern Approach*, 9th ed. New York: W. W. Norton & Company.