Travelling Wave Dynamics in Housing Markets

A Data-Driven Ricardian Explanation of Housing Market Crashes

Christopher Beech[1]

Abstract

We analyse England and Wales transactional house price data to describe a robust, empirical regularity across two cycles: a systematic progression of ranked annual returns. We demonstrate that whilst a geographical analysis by Local Authority appears intuitive, an analysis by quantile provides a superior description of information propagation through housing markets.

This travelling wave pattern, independent of the underlying house price growth, is both a data-driven version of Ricardo's Law of Rent and a leading indicator of house market crashes.

We propose a simple model where the log-logistic distribution of house prices is subject to drift and diffusion. In addition, a constant-speed travelling wave ensures that over time the fastest-growing quantile travels from the highest to the lowest in order. When the travelling wave reaches the lowest quantile or margin, there is a crash in house prices. The trough is reached when the land value at the margin is again worth zero; subsequently, the cycle restarts.

The benefits of a land value tax are briefly discussed.

The analysis makes testable statements not only about the future but also about other countries' national housing datasets.

Ricardo's Law of Rent

The classical economists Adam Smith, Thomas Malthus, and David Ricardo wrote extensively about land and the theory of rent[2]. Of particular interest is what has come to be known as Ricardo's Law of Rent, developed in Chapter 2 'On Rent' of 'On the Principles of Political Economy and Taxation'[3]. Ricardo's eloquent exposition of his idea is as follows:

> *If all land had the same properties, if it were boundless in quantity, and uniform in quality, no charge could be made for its use, unless where it possessed peculiar advantages of situation. It is only then because land is of different qualities with respect to its productive powers, and because in the progress of population, land of an inferior quality, or less advantageously situated, is called into cultivation, that rent is ever paid for the use of it. When, in the progress of society, land of the second degree of fertility is taken into cultivation, rent immediately commences on that of the first quality, and the amount of that rent will depend on the difference in the quality of these two portions of land.*
>
> *When land of the third quality is taken into cultivation, rent immediately commences on the second, and it is regulated as before, by the difference in their productive powers. At the same time, the rent of the first quality will rise, for that must always be above the rent of the second, by the difference between the produce which they yield with a given quantity of capital and labour. With every step in the progress of population, which shall oblige a country to have recourse to land of a worse quality, to enable it to raise its supply of food, rent, on all the more fertile land, will rise.*

There are two ideas here: the value of land, and how it changes over time.

Land's locational value in today's economy is equivalent to land's productive value in Ricardo's economy. A Georgian house in Liverpool is not worth as much as an identical property in London, and this difference is not due to the price of its materials but rather its land, or locational, value. (Labour costs will be higher in London, but this is a consequence of higher land values.) This is true of all houses, and there exists a continuum from where land consists of the majority of the value, to the margin where the value of land is nonexistent.

With respect to the change over time, it has been estimated that “Rising land prices explain about 80 percent of the global house price boom that has taken place since World War II.”[4]

Whereas Ricardo talks about ‘progress of population’ and the ‘progress of society’ and the change in rent that results, we analyse the change in median house prices over time. At each step in time, we rank the growth rate of all degrees of house prices. House prices are used as a proxy for land.

With the founding of neo-classical economics, land changed from being a third factor of production alongside capital and labour to being subsumed into capital[5]. As such, Ricardo and land are never mentioned, even when to do so would provide additional insight. For example, “In Britain, house prices exhibit a distinct spatial pattern over time, rising first in a cyclical upswing in the south-east and, then, spreading out over the rest of the country. This is known as the ripple effect.”[6]. Despite extensive literature reviews of the ripple effect in that paper and others[7,8] there is no mention of Ricardo’s Law of Rent.

Similarly, despite everything that has been written about housing markets since 2008, we have found no mention of Ricardo.

Land and housing cycles

Whilst mainstream models have largely ignored the role of land, a distinct strand of historical analysis has documented persistent periodicities in land values.

Hoyt[9] provided an empirical study of indexed land values in Chicago, recording peaks in 1836, 1855, 1872, 1890, 1910-12, and 1927. These intervals suggest a recurrence of approximately 18 to 20 years. This periodicity was subsequently formalised by Foldvary[10] within a Georgist-Austrian synthesis, which identified a regular 18-year cycle across the wider United States economy. Building on the work of Hoyt, Harrison[11] has also documented and predicted the last two 18-year house price cycles in the UK, where there have been house price peaks in 1973, 1989, and 2008.

That an approximate 18-year cycle has appeared across time and place in very different macroeconomic circumstances suggests an underlying mechanism not dependent on interest rates or credit conditions. Rather, we propose that this temporal periodicity is intrinsic to the propagation of information in housing markets.

Data analysis methodology

England and Wales house prices from 1995 onwards are available from the HM Land Registry Price Paid Data[12]. We exclude transactions representing the 'Other' category, leaving records for detached, semi-detached and terraced properties, and flats/maisonettes. We also exclude the Additional Price Paid category (just over 3% of transactions), as we do not know the additional price.

This data has then been matched to a Local Authority code using the ONS Postcode Directory[13], and then to a Local Authority District[14]. A mapping created by the author, of County and District to Local Authority, is used for those records without a Postcode or where there is no match in the ONS PD. The Isles of Scilly are excluded as there are too few sales. Stamp Duty is calculated and included. We exclude data beyond September 2025 due to a lag in registrations at Land Registry.

Given the wide range of house prices and compositional change each month per Local Authority, the median is chosen as the best measure of central tendency. We order the Local Authorities from high to low using the 2025 median price. The annual log return of the median is calculated per Local Authority, and a 2x12-month moving average smooths out seasonality effects. Partial computations are allowed, as such the moving average is less robust at the beginning and end of the time series. A rank of the growth rates is derived from the moving average; we refer to 'rank […] of the return', as the rank is the same for the return as it is for the log return.

This analysis is then repeated using a price per square metre dataset created by matching house prices to EPC data, which includes the size of the property. Our code[15] is a refactoring for speed of the work by Chi et al.[16].

An analysis by quantile[17] is then conducted for both datasets instead of by Local Authority; 317 quantiles are used to match the number of Local Authorities. Unlike the Local Authorities, the quantiles are recomputed per month. Quantiles have the benefit of avoiding both the heterogeneity of prices within a Local Authority and arbitrary administrative borders.

Heat maps are created of the median price per Local Authority or quantile over time. This approach is based on a similar heat map[18] of Local Authorities from 2017 by Neal Hudson of consultancy Built Place. Formally, this is a change of the coordinates of the system from the usual analysis of house price indices.

Whereas Hudson's heat map also shows an overlay of the single fastest-growing Local Authority over time, we created a separate plot showing all ranks for Local Authorities or quantiles. We then group by deciles to reduce the noise in the plots.

Adjacency matrices are created from spatiotemporal networks of the ranks. The spatial network captures the influence of a Local Authority or quantile on another at the same time step, that is, rank one influences rank two and so on. The temporal network captures the influence of a Local Authority or quantile on another at consecutive time steps, that is, rank one influences rank one at the next time step and so on. By definition, only the temporal network can capture edges between the same Local Authority or quantile. These edges are combined and weighted, a graph created, and the adjacency matrices calculated from the graph.

The Supplementary Material provides additional analyses and plots of: the probability density and empirical cumulative distribution functions; Cullen and Frey per year and per Local Authority; time series decomposition heat maps; the mean and standard deviation of the log returns; a Principal Component Analysis, K-means clustering of the log returns; credit versus median prices; repeat sales; and a price-to-earnings ratio per Local Authority.

Results

Figure 1 shows the moving average of the annual log return of house prices. The top row is by Local Authority, the bottom row by quantile. The left column is calculated from price, the right column from price per square metre.

In the cycle to 2008, median house prices initially increased in the highest-priced areas, then over time increased across the middle and lower-priced areas. The crash is very clear in blue across all Local Authorities. In the bottom left quantile plot, we see a stepped pattern due to the old Stamp Duty regime.

From 2010, there is the start of a recovery before a modest fall in prices in the austerity years. The pattern starts again afterwards but without such high log returns as previously. Later there is the impact of the pandemic and the subsequent increases in property prices before the aftermath of the 2022 mini-budget.

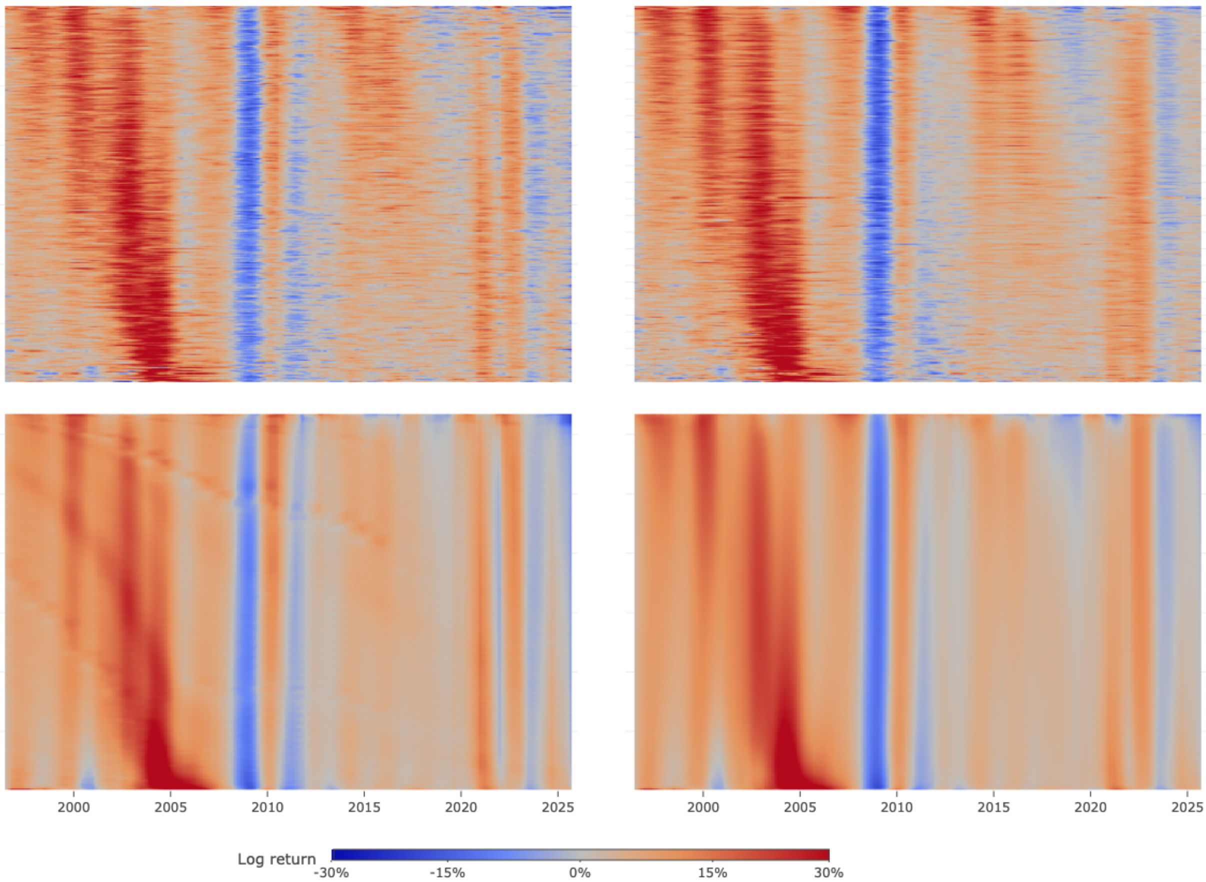

Figure 1: Moving average of the annual return of the monthly median; highest median price at the top, lowest median price at the bottom. By Local Authority (top left), by Local Authority using the price per square metre dataset (top right), by quantile (bottom left), and by quantile using the price per square metre dataset (bottom right).

Figure 2 shows the rank of the returns over time, allowing us to visualise the change in house prices independently of the underlying house price growth. The top row is by Local Authority, the bottom row by quantile. The left column is calculated from price, the right column from price per square metre. The fastest growing (or lowest falling) areas are in red, and the slowest growing (or highest falling) areas are in blue.

In the bottom left quantile plot, we once again see a stepped pattern due to the old Stamp Duty regime.

In all plots, there are clear slanted red lines representing the fastest-growing areas in the house price cycle to 2008 and in the cycle from 2010. The rate of progress is very similar in both cases. These patterns we define as travelling waves.

It is noticeable that the travelling wave reaches the Local Authority or quantile with the lowest median price prior to 2008. However, the richest areas subsequently saw the fastest growth immediately prior. This may have been due to foreign buyers in London, although we can find no evidence. Alternatively, it may be what Harrison[19] calls a 'winner's curse' phase.

In the second cycle from 2010, the travelling wave starts again before briefly stalling in the austerity years. When it does continue beyond the upper Local Authorities or quantiles, it is parallel to the pattern from the earlier cycle. There is no impact on the pattern from Brexit. However, it is interrupted due to the pandemic when upper middle Local Authorities and upper quantiles saw the fastest growth, due to the so-called 'race for space'[20].

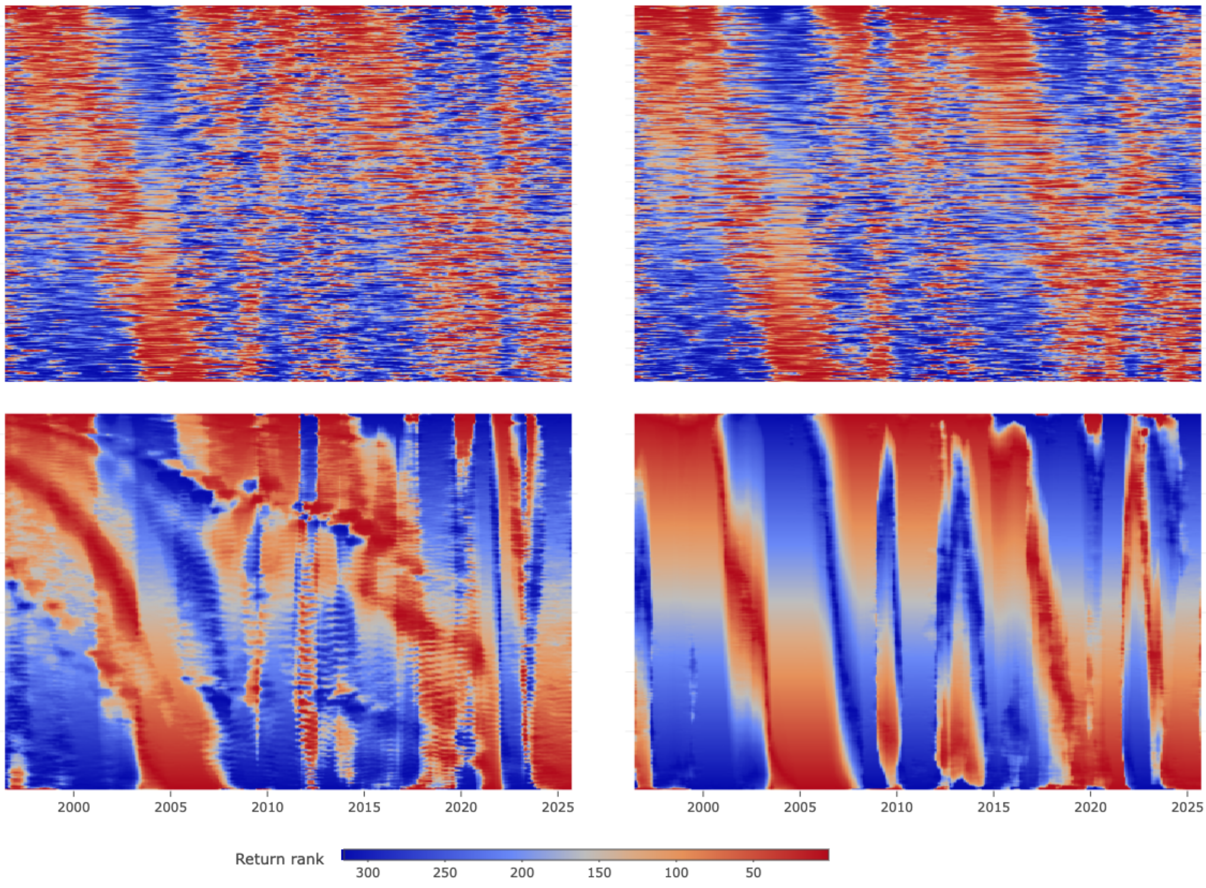

Figure 2: Rank of the moving average of the annual return of the monthly median; highest median price at the top, lowest median price at the bottom. By Local Authority (top left), by Local Authority using the price per square metre dataset (top right), by quantile (bottom left), and by quantile using the price per square metre dataset (bottom right).

Figure 3 repeats Figure 2 but using deciles. This reduces the noise in the data. It is striking that in the quantile price per square metre plot (bottom right), the travelling wave ranks are particularly clean (systematically ordered from ranks one to ten) in both vertical and horizontal dimensions.

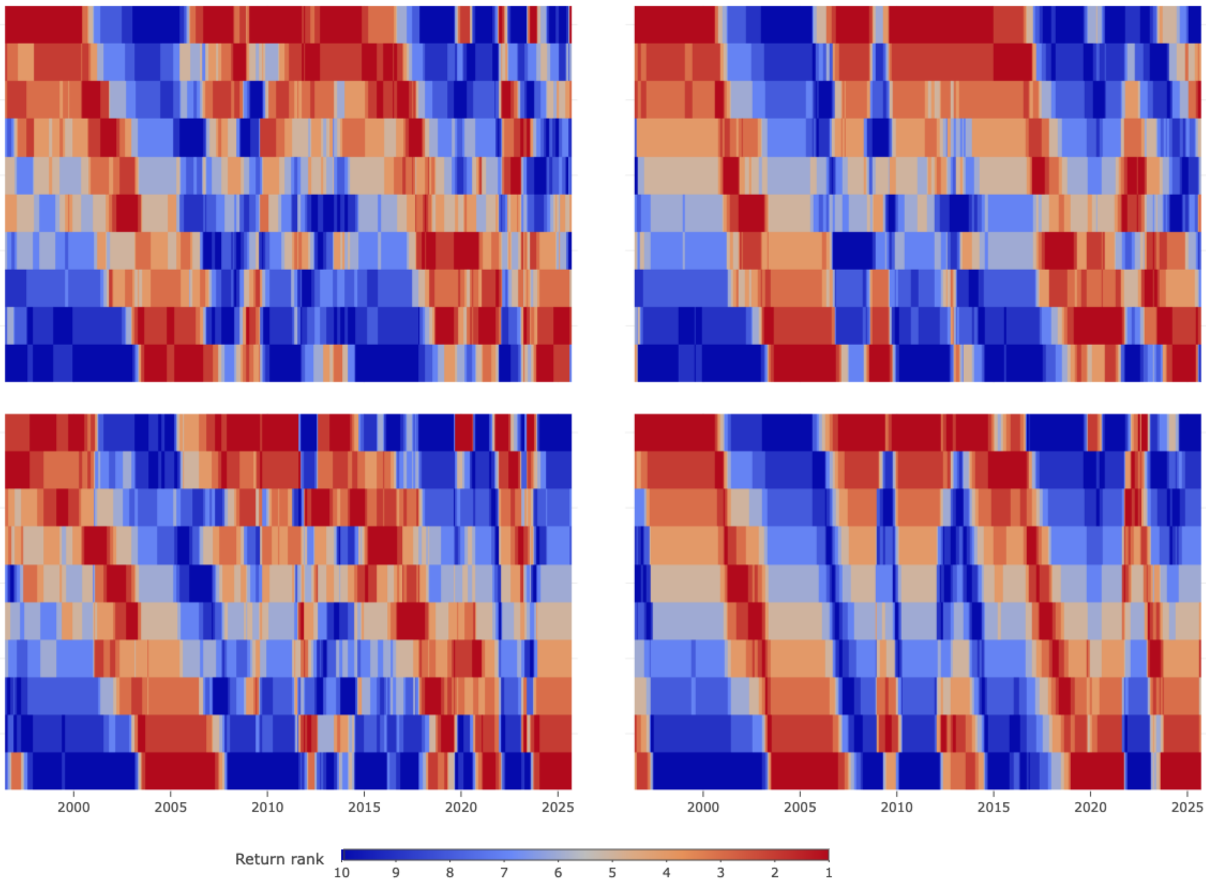

Figure 3: Rank of the moving average of the annual return of the monthly median, by decile; highest median price at the top, lowest median price at the bottom. By Local Authority (top left), by Local Authority using the price per square metre dataset (top right), by quantile (bottom left), and by quantile using the price per square metre dataset (bottom right).

Figure 4 combines both the log return and rank decile plots for comparison.

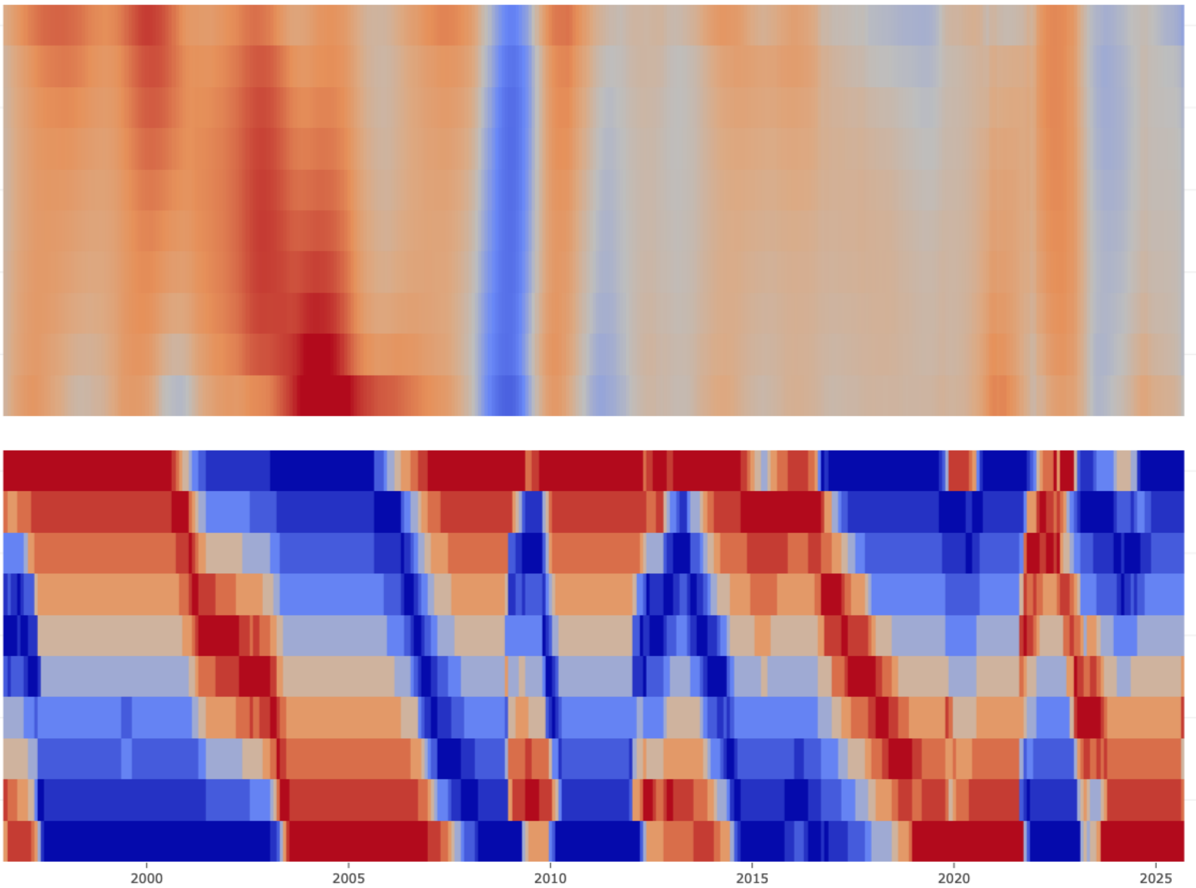

Figure 4: Top plot: Moving average of the annual log return of the monthly median using the price per square metre dataset, by decile. Bottom plot: Rank of the moving average of the annual return of the monthly median using the price per square metre dataset, by decile. Highest median price at the top, lowest median price at the bottom.

Figure 5 shows the adjacency matrices.

There is a weak adjacency pattern for all Local Authorities (top row). This weak adjacency pattern is expected, given the heterogeneity of property values within Local Authorities, the sold prices at any given point in time are only a sample of the underlying distribution, and the mix will change, maybe significantly, from one period to the next.

However, there is a strong adjacency pattern for the quantiles (bottom row), even more so for the price per square metre dataset (bottom right).

Furthermore, if it were possible to separate the land value from the overall house price (at any given quantile there will be a group of house prices with different mixes of land and materials), then the adjacency pattern would be stronger still.

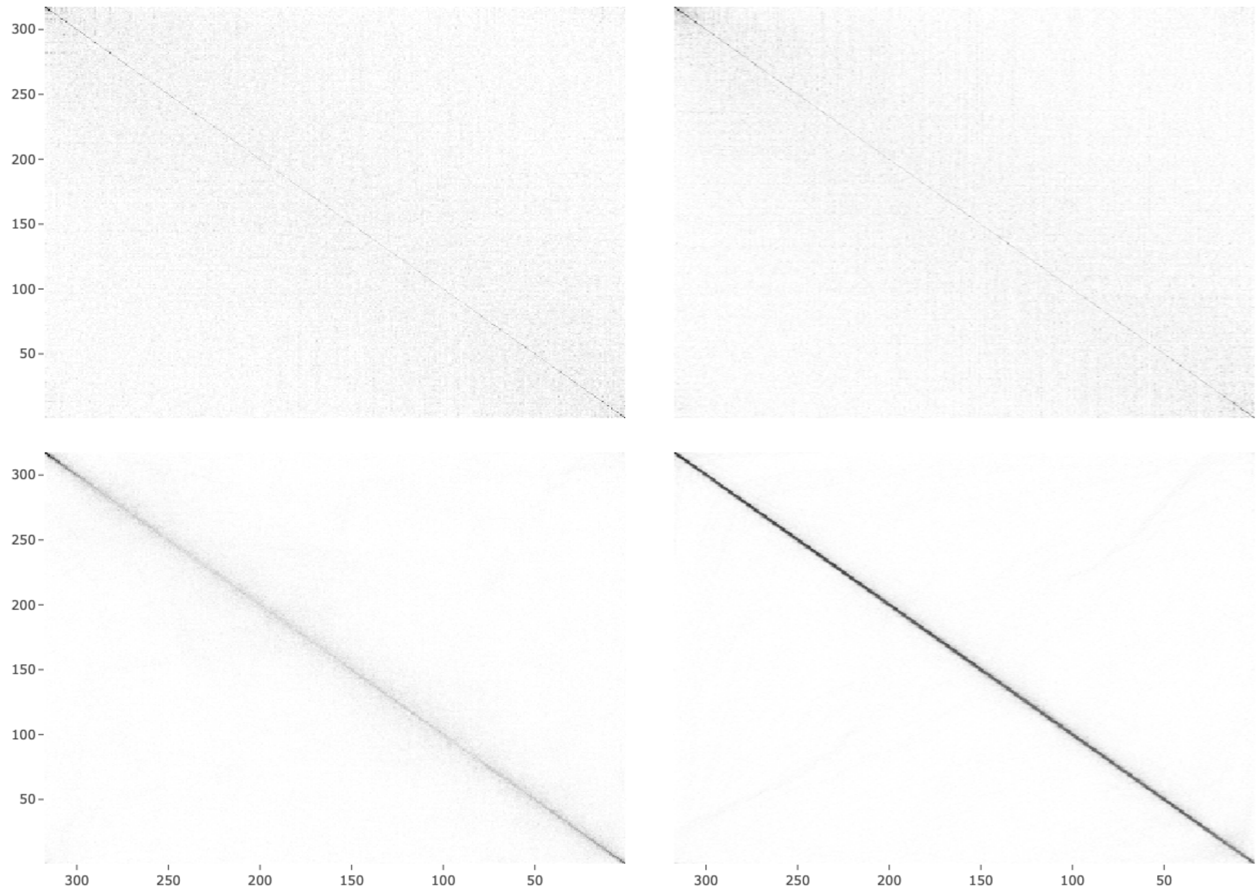

Figure 5: Adjacency matrices of the ranks; highest median price at the top, lowest median price at the bottom. By Local Authority (top left), by Local Authority using the price per square metre dataset (top right), by quantile (bottom left), and by quantile using the price per square metre dataset (bottom right).

Discussion

By ranking the moving average of the annual return of the monthly median, a distinct travelling wave pattern has emerged for two house price cycles. This is most clearly seen using price per square metre data per quantile.

Could the travelling wave be an early warning signal[21], a leading indicator, for the end of a house price cycle? And why should housing cycles end once the lowest quantile is the fastest growing?

Markets are, of course, producers and consumers of price information.

Where Ricardo talks about land of different degrees of fertility being taken into cultivation in turn, there are direct parallels with the travelling wave, in both the rank and adjacency matrix plots.

The travelling wave follows the reversed sigmoid of the empirical cumulative distribution function of log house prices from high to low. This suggests a constant speed for the travelling wave. Although this speed is the rate of information transfer, we do not know why it is this speed.

The travelling wave represents the market's dynamic adjustment process as it seeks to maintain a no-arbitrage condition across quantiles, and this is clearly seen in the adjacency matrices: no gaps in the distribution occur as prices rise or fall; each quantile influences its adjacent quantiles. Despite housing markets being subject to crashes, in this respect, they are efficient.

For the highest-priced house, the majority of the value will be the land. As prices decrease, the value of the land is, as a general rule, lower relative to the cost of the materials. At the margin, there will be a house (or set of streets) where the only value is the rebuild cost; the productive value of the land is zero.

In the classical view of Ricardo and Malthus, extending the margin of cultivation forces living standards down to subsistence levels. In our model, a similar boundary is reached when the travelling wave hits the lowest quantile. At this point, the market price of the marginal unit far exceeds its rebuild cost, implying a positive land value, whereas Ricardian theory dictates it should be zero. There is a collective realisation that the market value is incorrect, colloquially: there is no distance left to run. The market crashes to restore the marginal land value to zero.

To further support this explanation, the repeat sales analysis shows that the trough of the crash is indeed when houses at the margin are again worth their rebuild cost. Once this happens, the cycle can start over.

Whilst Ricardo provides an economic basis for our discussion, causal inference is more difficult.

To model causation, we hypothesise that we could use a transfer entropy approach building on information theory: a housing cycle will end given that the travelling wave has progressed through the full range of the probability density function.

However, even if the appropriate formulation were readily available, we see two major exceptions to this in the real-world data: i) immediately prior to 2008, it was the highest quantiles that were the fastest growing, and ii) as a result of the pandemic, the travelling wave was interrupted before resuming its regular pattern.

Approaches to modelling

Many models of house prices exclude the possibility of bubbles since they are based on a simple SDE with standard Brownian motion[22]. This implies that the distribution of prices is log-normal.

The system could be modelled as an n-dimensional set of SDEs, but the weak adjacency pattern for the Local Authorities excludes this possibility. Additionally, although some Local Authority log return distributions are normal and could be modelled using a closed-form solution[23], many are not, and no closed-form solution exists.

It is best to consider the time evolution of the overall probability density function, as in simple SDE models, but with a small number of changes and additions:

The log-logistic distribution of house prices is subject to drift and diffusion. A constant-speed travelling wave ensures that over time the fastest-growing quantile travels from the highest to the lowest in order. When the travelling wave reaches the lowest quantile, the boundary or margin, there is a crash in house prices[24].

The trough is reached when the land value at the margin is again worth zero; subsequently, the cycle restarts.

A normative extension to the model: Land Value Tax

Given the inevitability housing cycle ending, how can the impact be avoided, or at least minimised? The lower the growth in house prices over a cycle, the less damage to the economy.

The simplest mathematical approach would be to include some stabilising negative feedback: a land value tax[25,26,27]. This would affect the drift directly, the diffusion indirectly, but not the travelling wave.

By offsetting a land value tax against income taxes, the house price-to-earnings ratio would be reduced, and more households could afford to buy.

A land value tax would also increase the supply of land for development by reducing its option value[28]. Theoretically, “Per the Hotelling rule, the (unimproved) value of land should increase at the real rate of interest.”[29] Since the value of land grows at a greater rate than this, there is no incentive to develop; the sensible option is to hold and do nothing.

Conclusion

Our analysis demonstrates that ranking the annual returns of housing transactions reveals a distinct travelling wave pattern across two market cycles. This pattern is most clearly observed when shifting from a geographical analysis by Local Authority to a value-based analysis by quantile. The travelling wave is independent of the underlying house price growth, suggesting it is an intrinsic property of market structure.

Furthermore, our analysis of spatiotemporal adjacency matrices clarifies the mechanism of information transmission within the market. While the influence between Local Authorities is shown to be diffuse and weak, the influence between price quantiles is strong and systematic. This finding indicates that the housing market satisfies a no-arbitrage condition across the price distribution; information and price adjustments propagate sequentially through adjacent quantiles.

We argue that the travelling wave is both a data-driven version of Ricardo’s Law of Rent and a leading indicator of house price crashes[30]. When combined with the 18-year cycle previously observed by others, the current location of the travelling wave suggests a crash in 2026-27.

We hypothesise that this methodological framework is not specific to England and Wales and will reveal similar structural dynamics in other countries’ national housing datasets[31].

Supplementary Material

A. Neal Hudson's heat map

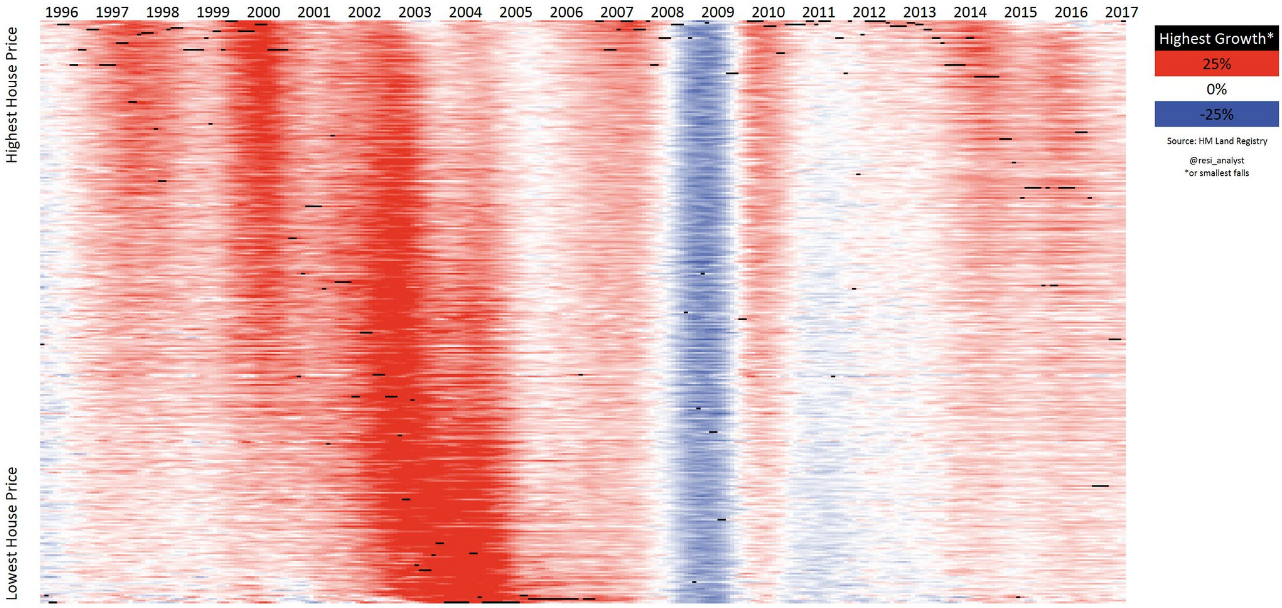

Figure SM.1: Annual change in Local Authority house prices, by Neal Hudson of Built Place https://builtplace.com

Ordering of Local Authorities was based on the latest price; growth was based on a three-month rolling average of monthly prices. The growth did not use a log.

The overlay of black dots shows the single fastest-growing Local Authority at any point in time.

B. Guide to the R scripts and additional analysis

1. Import and tidy

This code is described in the main text above. With minor modifications, it is also used to prepare data for the repeat sales analysis.

The price per square metre dataset is also imported and prepared.

2. Stamp Duty

There is often confusion about Stamp Duty in the popular press with the idea that a cut in Stamp Duty will not lead to an increase in house prices, but this is incorrect unless the seller is particularly naïve (although markets may take time to adjust). As RICS guidance explains, 'land is the residual'[32]; the supply of land is inelastic. Hence, the proper price of a property is the price including Stamp Duty. We are not able to determine whether a transaction is for a second home at a higher rate of Stamp Duty, so we apply the basic rates to all. Repeating the analysis without Stamp Duty makes no difference to the results, but we are thorough where possible.

Stamp Duty has been calculated using historical rates[33,34,35] and using the appropriate method: based on the highest band the price falls into (before December 3, 2014) or based on the portion within a given band (thereafter). Similarly, values for England and Wales, England alone, or Wales alone have been used, as appropriate.

3. Yearly distributions

In Figure SM.2, the Cullen and Frey plot shows that the probability density functions are log-logistic since the log price is logistic with only minor skew. The price per square metre dataset is not too dissimilar (not shown).

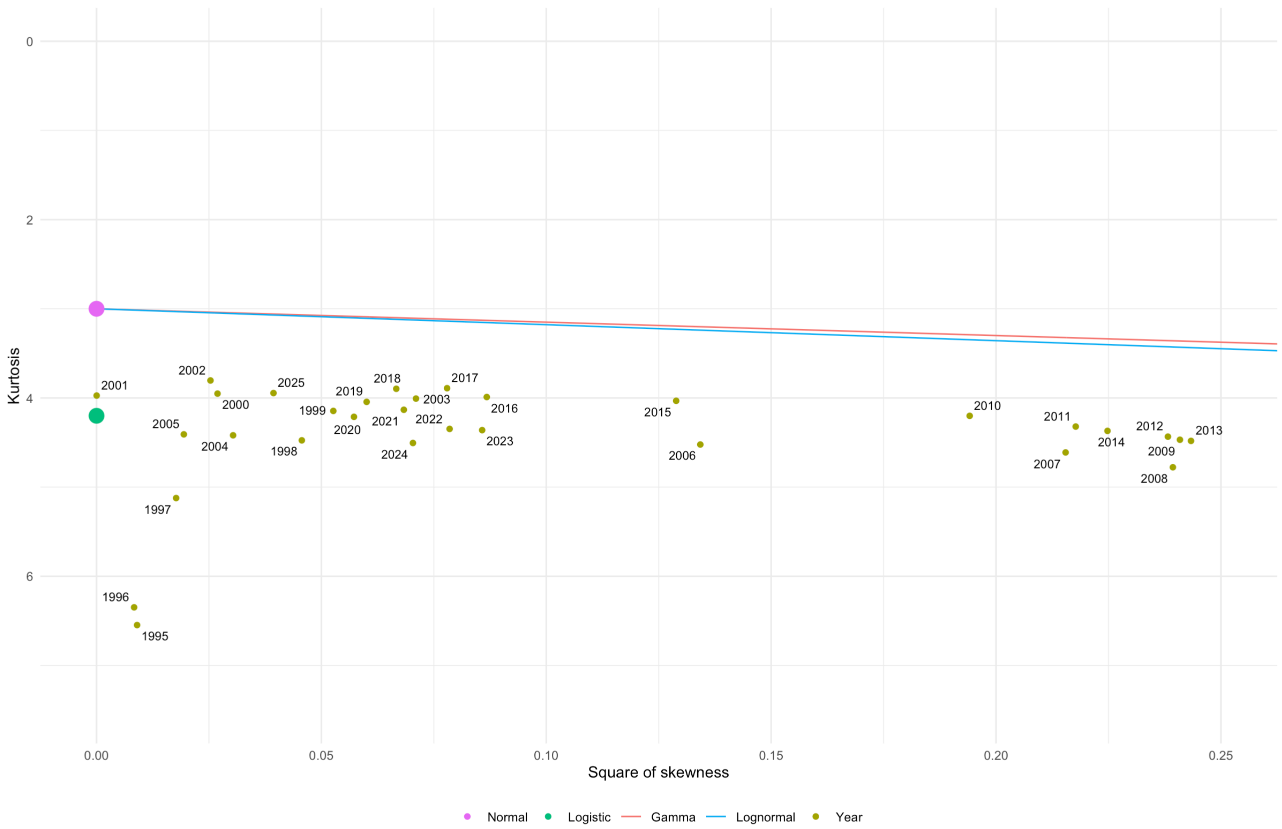

Figure SM.2: Log price kurtosis and square of skewness, by year

Figures SM.3 - SM.6 show the probability density functions. The effect of the old Stamp Duty regime is shown in the spikes in transaction numbers just below the thresholds.

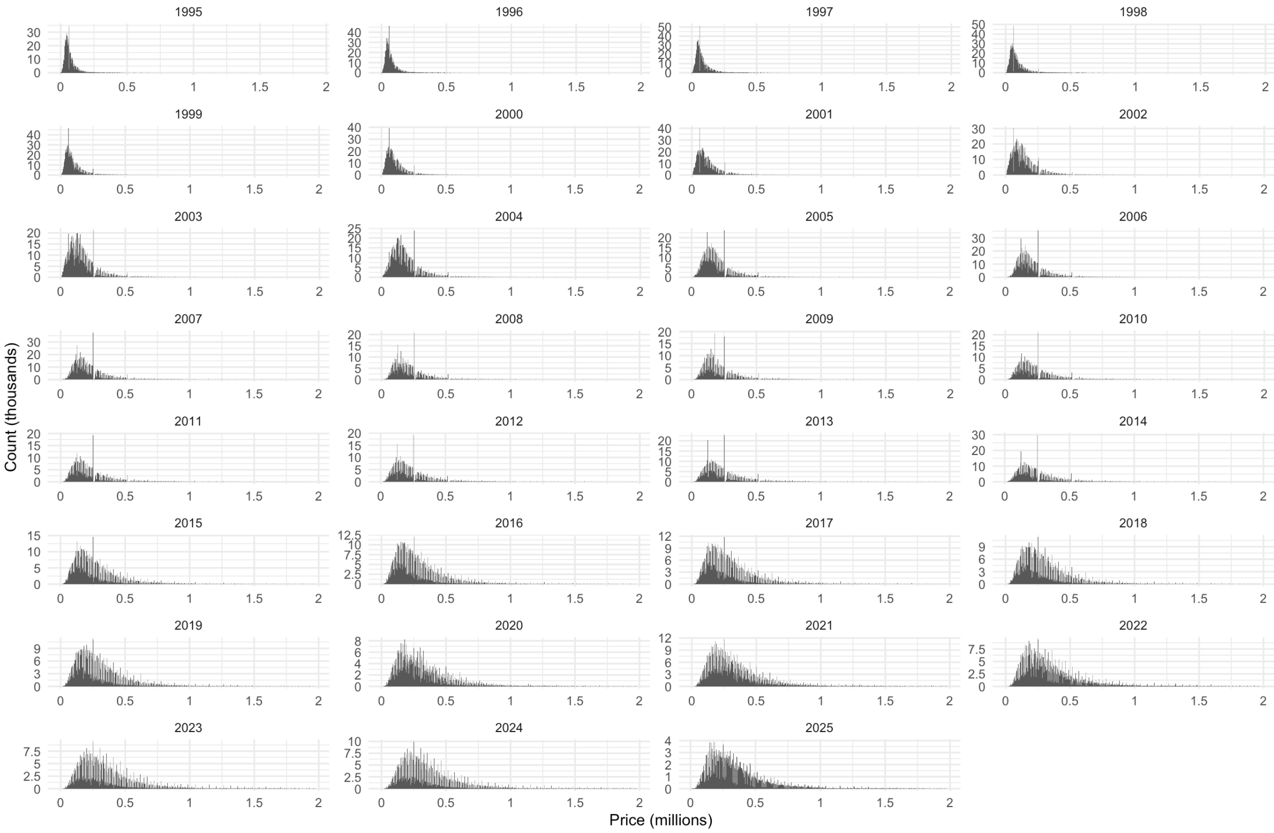

Figure SM.3: Price pdf (histogram), by year

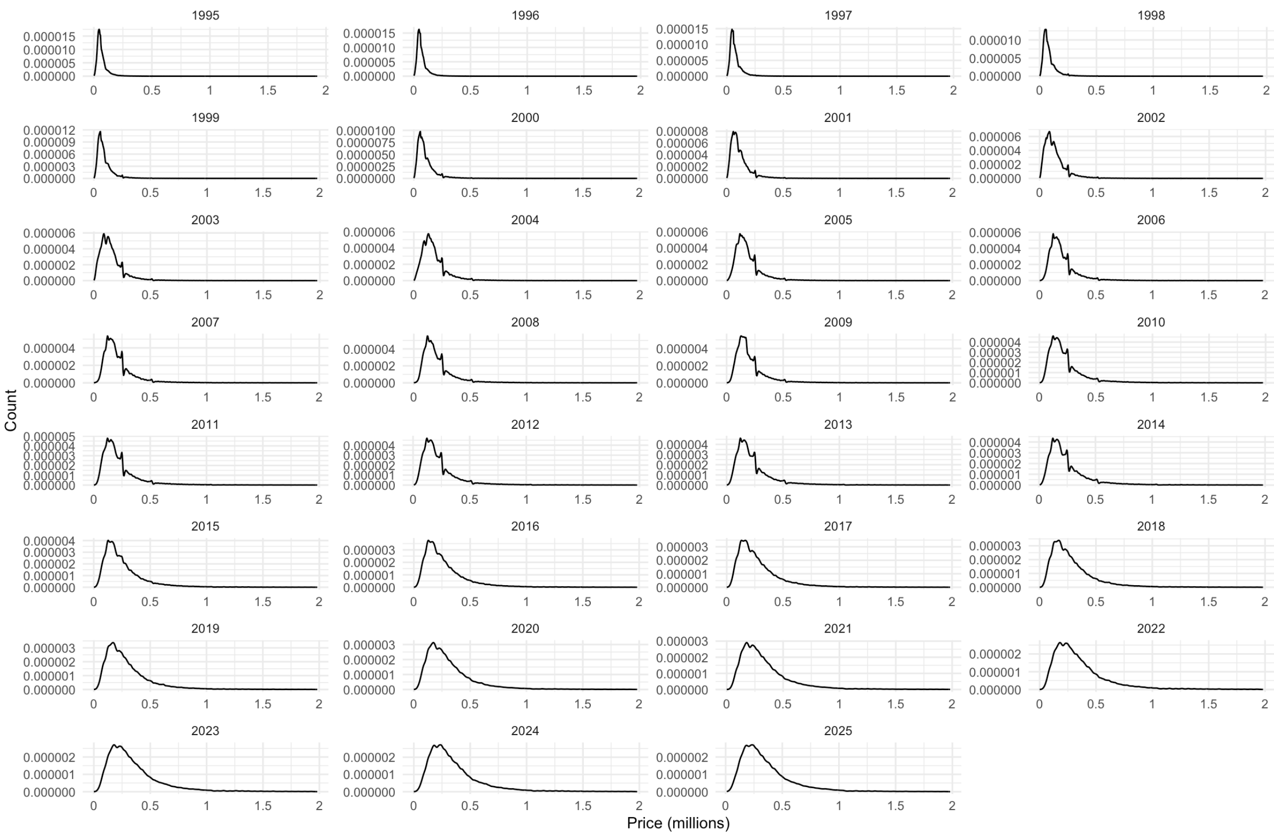

Figure SM.4: Price pdf (density), by year

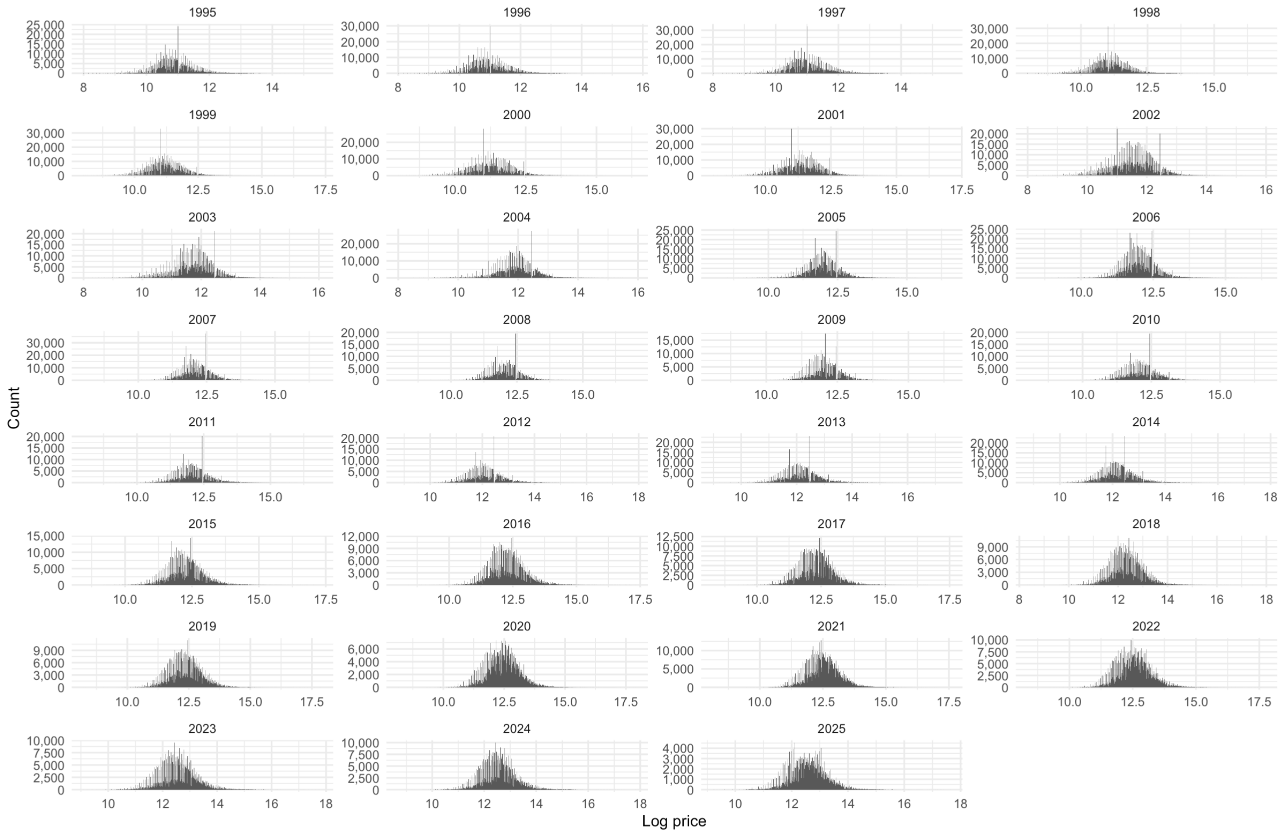

Figure SM.5: Log price pdf (histogram), by year

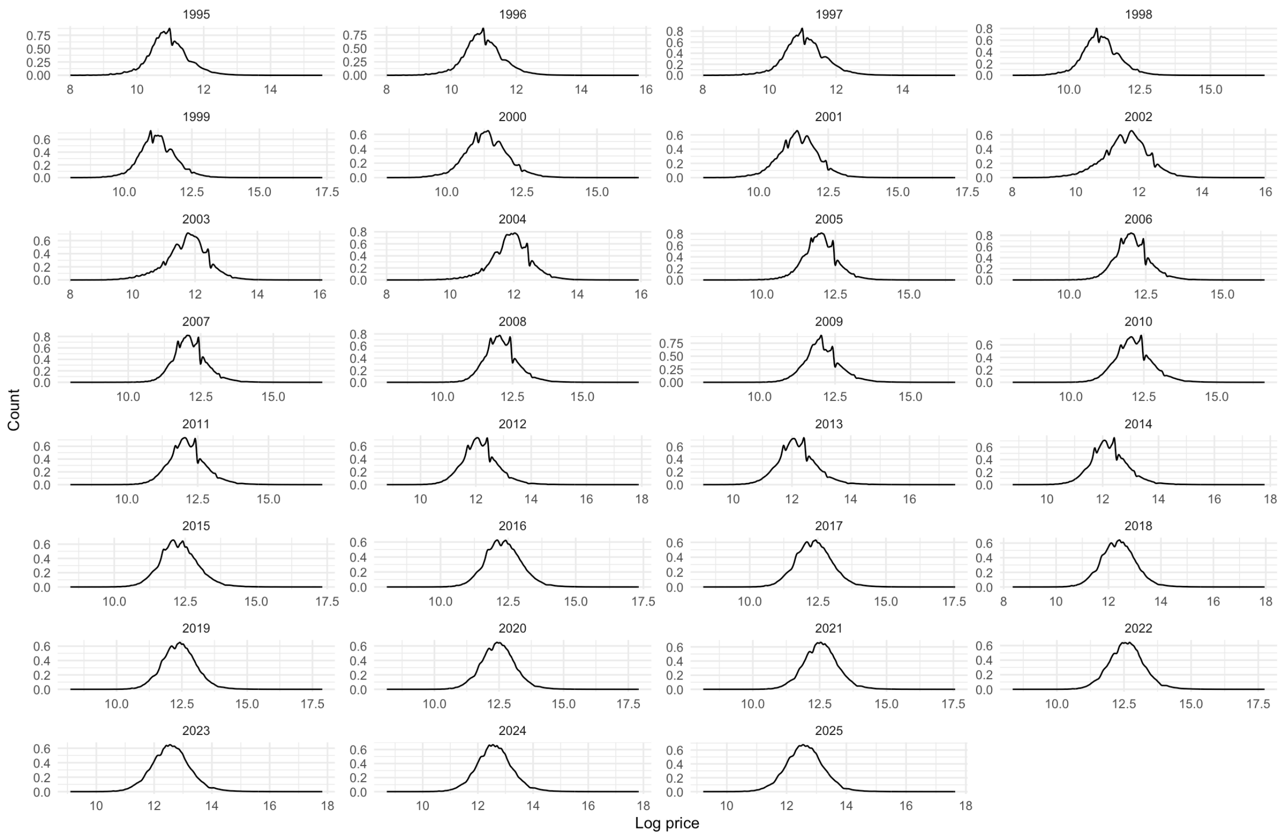

Figure SM.6: Log price pdf (density), by year

Figure SM.7 shows the empirical cumulative distribution function by year, a set of sigmoid curves.

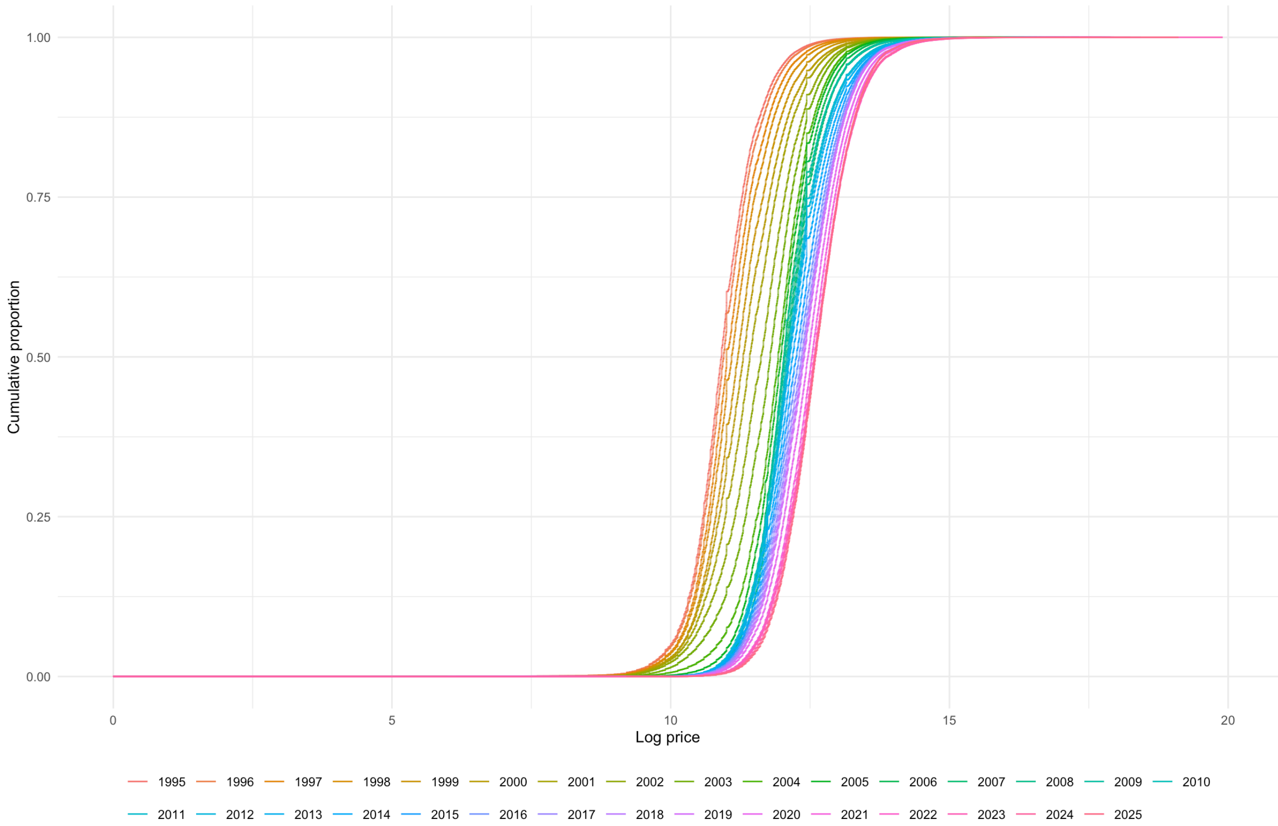

Figure SM.7: Empirical cumulative distribution function (cumulative log_price), by year

4. Local Authority distributions

Preliminary analysis is conducted as to the best way to calculate the central tendency of the distributions per Local Authority.

Are there enough sales in each month to allow a meaningful analysis? Figure SM.8 suggests so.

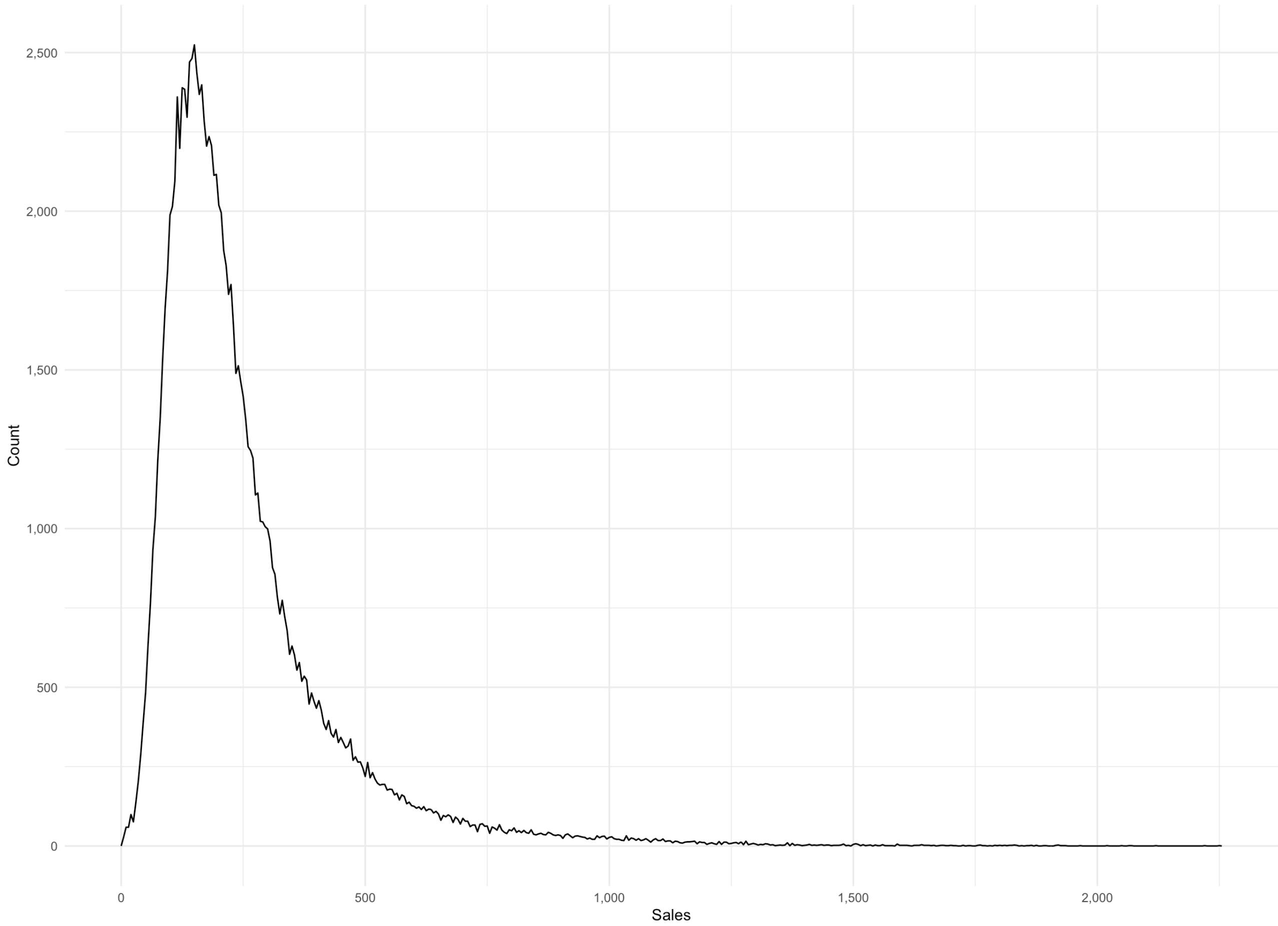

Figure SM.8: Sales frequency by month, all Local Authorities

It is possible to group transactions by quarter or year, but this results in very asymmetrical returns as the periods are not aligned with seasonality effects.

Lognormal and right Pareto lognormal distributions can be fitted, but the median is the only sensible and robust approach: Figure SM.B.9 shows such a wide range of distributions at a Local Authority level. Calculation of the median requires a little bit of noise added (a normal distribution with a mean of 0 and standard deviation of 1 is sufficient) to ensure that repeat values do not stop the median being calculated.

The price per square metre dataset shown in Figure SM.B.10 has far less kurtosis and skew, and this makes intuitive sense.

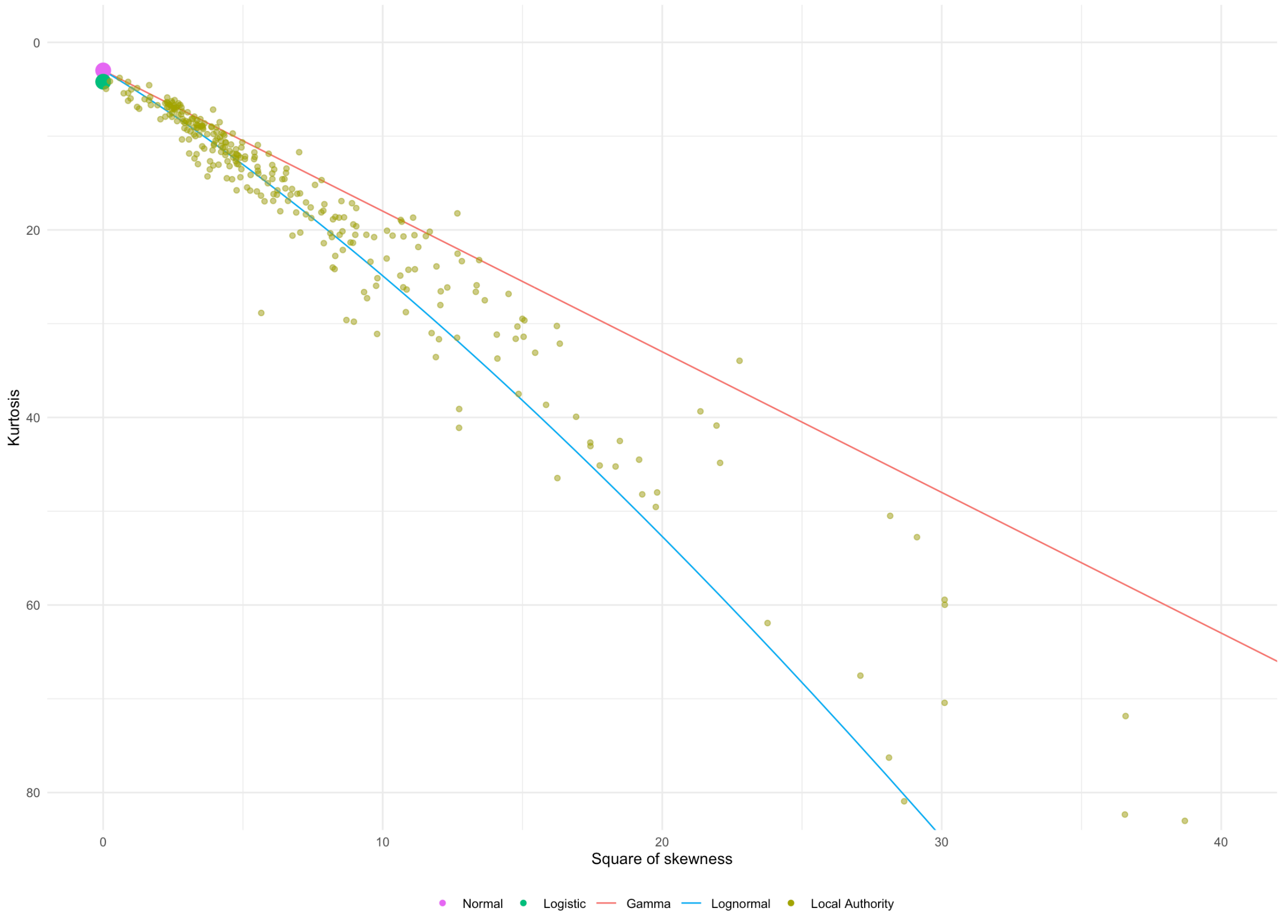

Figure SM.9: Price pdf kurtosis and square of skewness, by Local Authority (2025)

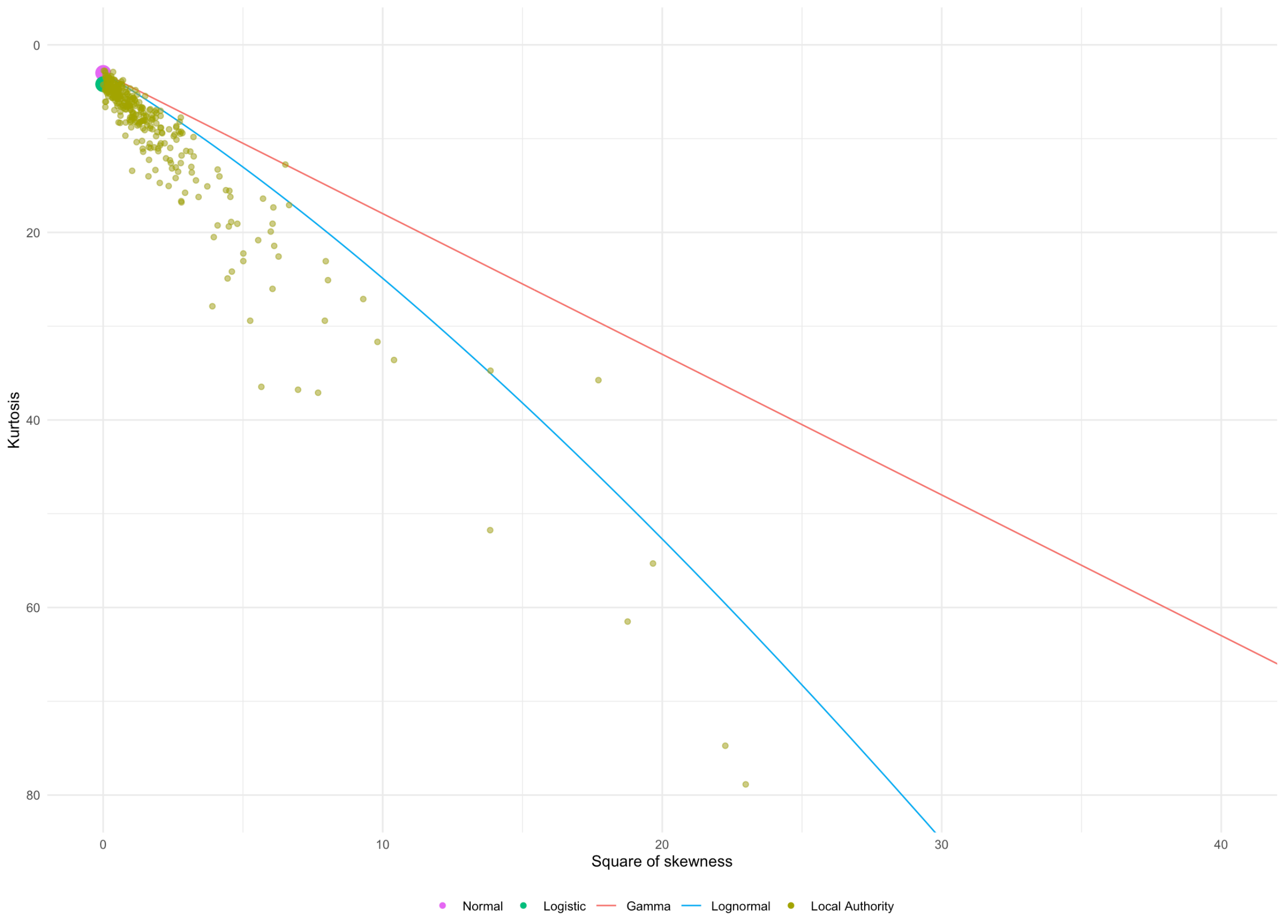

Figure SM.10: Price (per square metre) pdf kurtosis and square of skewness, by Local Authority (2025)

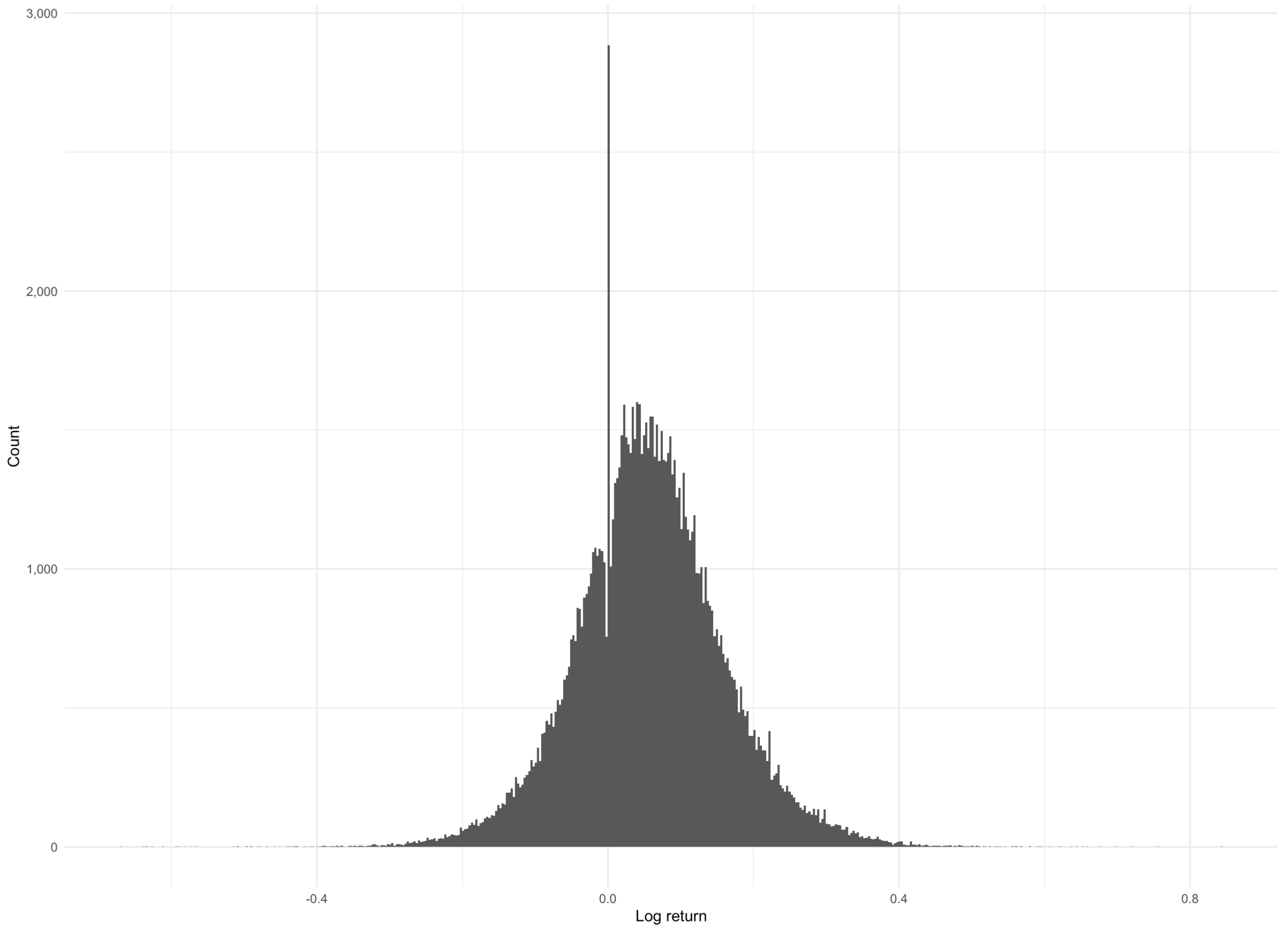

Figure SM.11: Annual log return of the median price, all Local Authorities

In Figure SM.11, the annual log return of the median price of all Local Authorities is logistic (an estimated kurtosis of 4.31 and an estimated skewness of 0.1). Of most interest is the spike and surrounding gaps in transactions at zero.

In Figure SM.12, the median price annual log return pdf by Local Authority varies between normal to logistic and surrounding values.

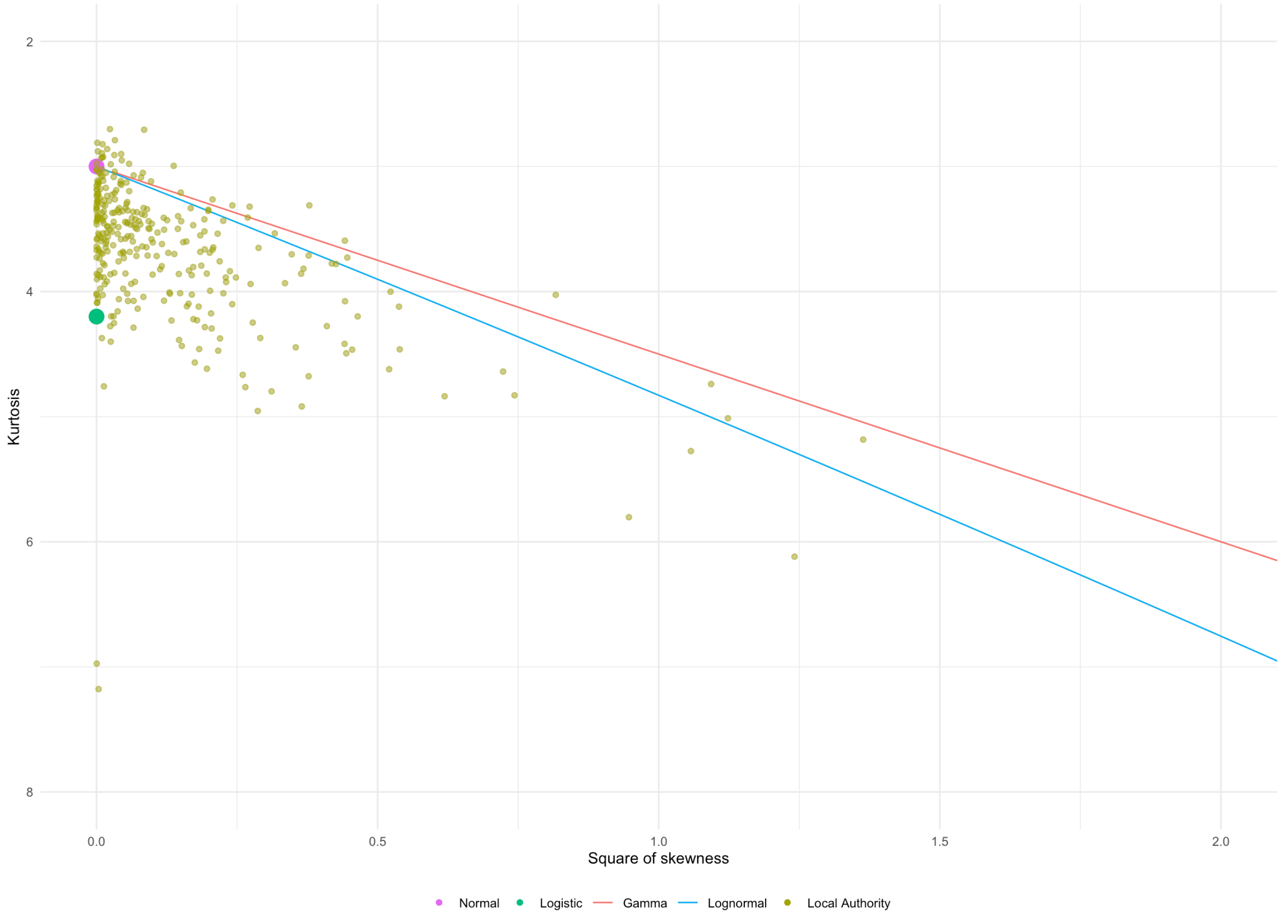

Figure SM.12: Median price annual log return pdf kurtosis and square of skewness, by Local Authority

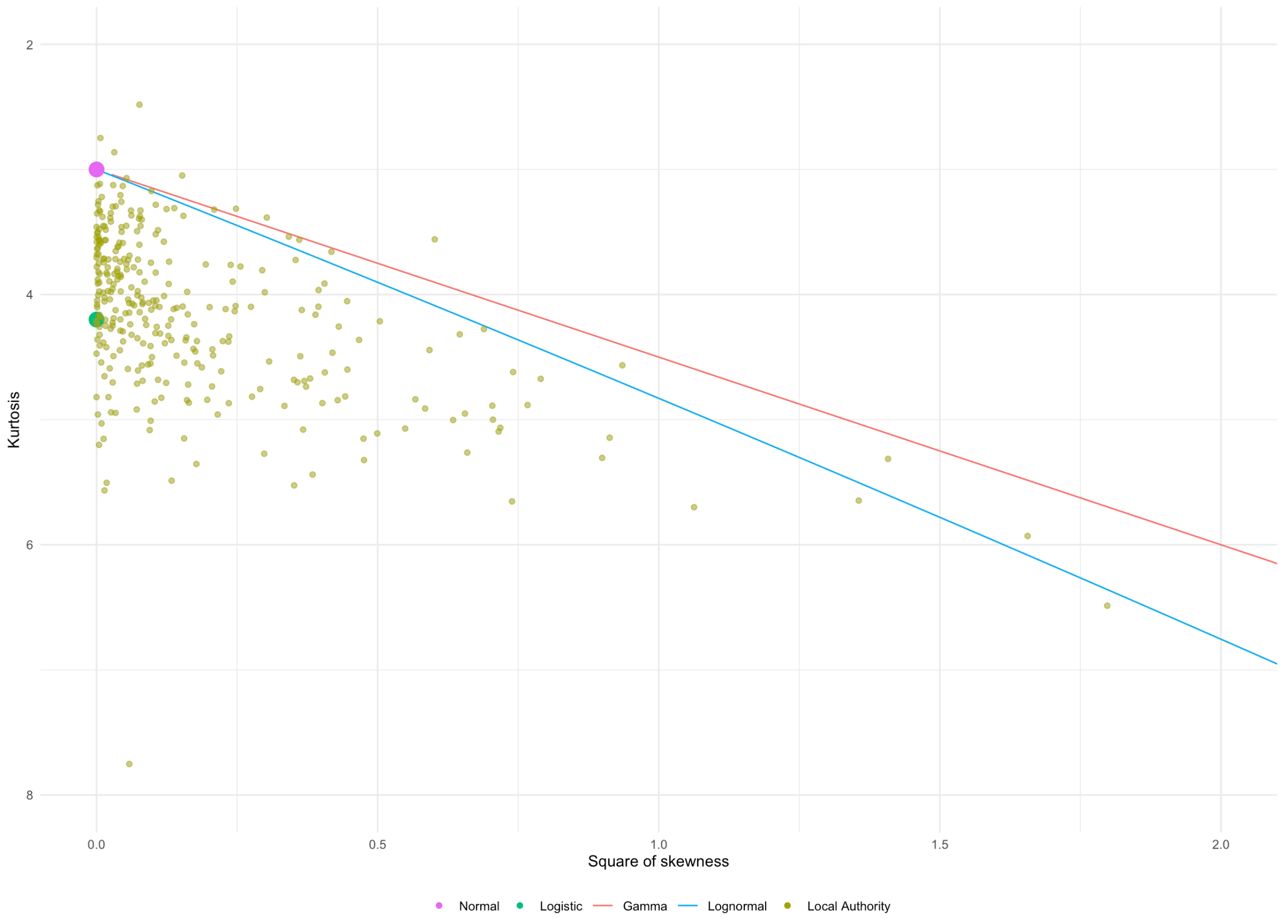

Figure SM.13: Median price (per square metre) annual log return pdf kurtosis and square of skewness, by Local Authority

5. Local Authority analysis

In Figure SM.14, the mean and standard deviation of the log returns are plotted for the entire time series. It should be no surprise, given the log-logistic pdf, that the Local Authorities with higher median prices also have higher growth rates.

Figure SM.14: Mean and standard deviation of the log return of the median price, by Local Authority; highest median price Local Authority at the top, lowest median price Local Authority at the bottom.

Figure SM.15 combines a Principal Component Analysis with a K-means clustering of the log return. From this alone, the north/south housing divide is derived. Note: the PCA axes have been changed to allow alignment with the convention of Local Authority on the y-axis, with the highest median at the top.

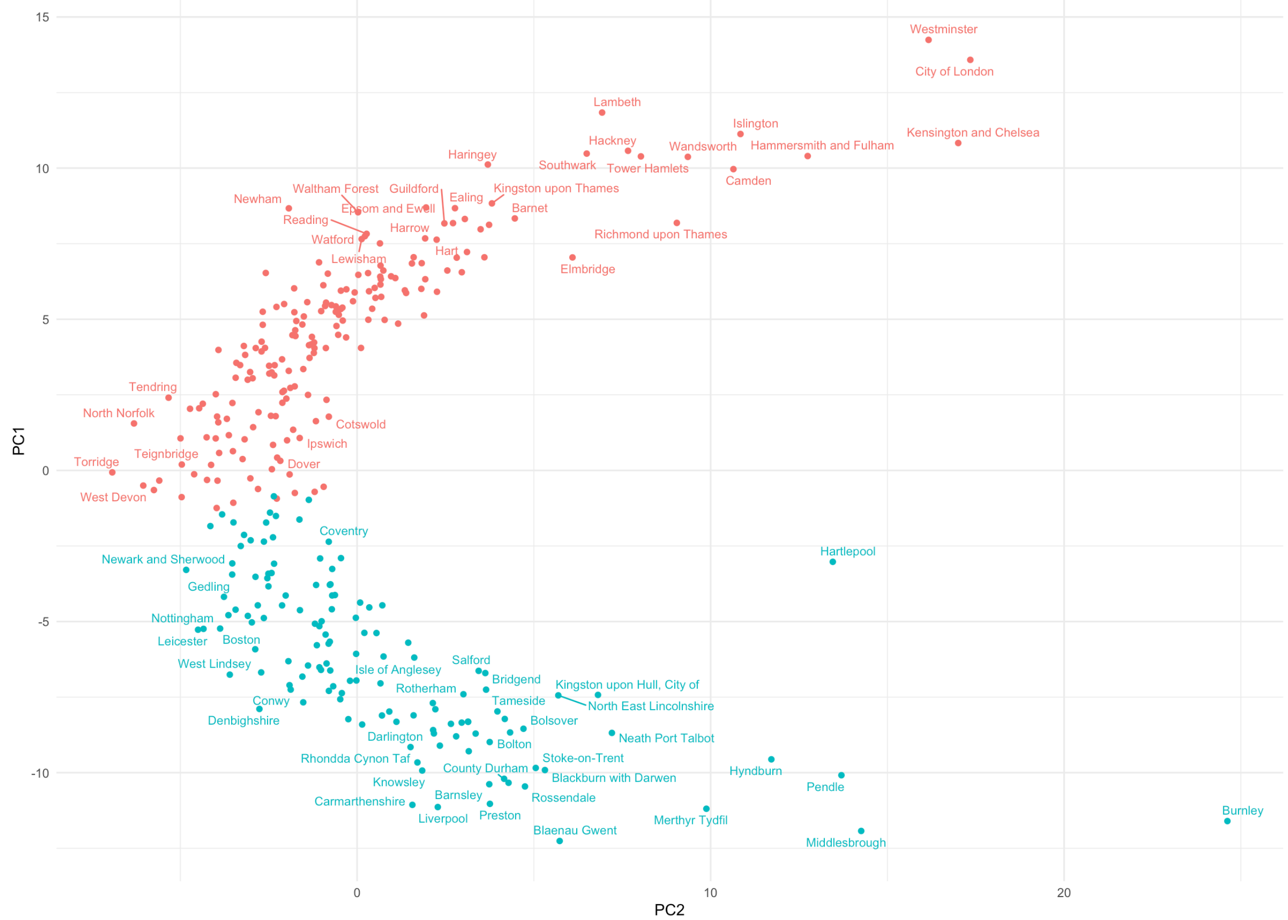

Figure SM.15: Principal Component Analysis, K-Means Clustering (n = 2), by Local Authority

And using three clusters, Figure SM.16 separates out central London, plus Richmond and Elmbridge.

Figure SM.16: Principal Component Analysis, K-Means Clustering (n = 3), by Local Authority

5a. Local Authority maps

Due to the amount of time it takes to render the ONS shape file[36], code for the maps has been moved into a separate script.

There are two maps:

- In Figure SM.17, a map of the (log) median price; the log has been taken else Kensington and Chelsea distorts the colour scale.

- In Figure SM.18, a map of the north/south divide, which is not that different from the work of Danny Dorling[37].

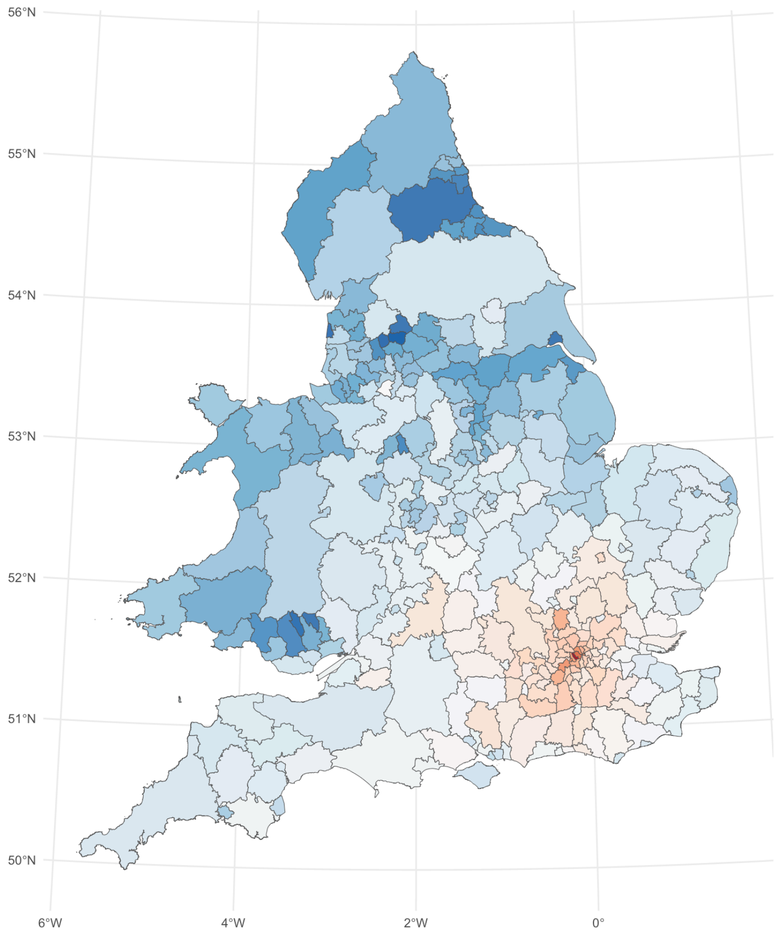

Figure SM.17: North South divide - log median price

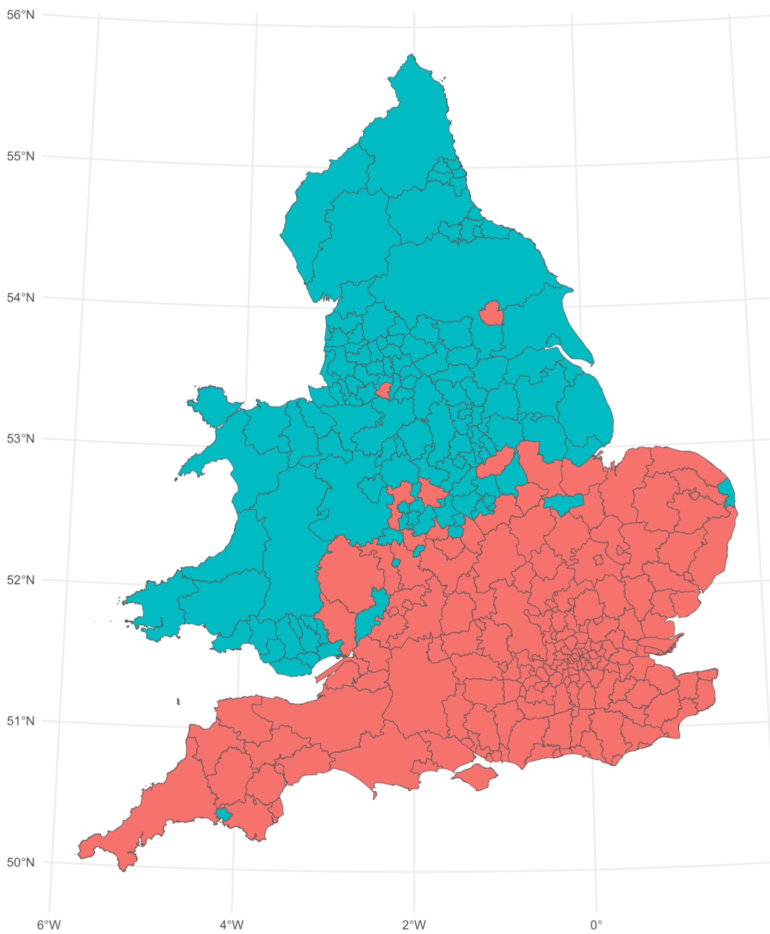

Figure SM.18: North South divide - K-Means clustering of log returns

The final piece of analysis is a spatiotemporal network graph of the rankings.

For the spatial edges, the medians are arranged (ordered) by Date, then Rank, then grouped by Date. The associated spatial edge is the next Local Authority in the list.

```
arrange(Date, Rank) |>
  group_by(Date) |>
  mutate(to = lead(`Local Authority`))
```

For the temporal edges, the medians are arranged (ordered) by Rank, then Date, then grouped by Rank. The associated temporal edge is the next Local Authority in the list.

```
arrange(Rank, Date) |>
  group_by(Rank) |>
  mutate(to = lead(`Local Authority`))
```

These edges are combined, and edge weights, a graph of these edge weights, and an adjacency matrix from the graph are calculated. This graph is plotted to show the relative influence of each Local Authority on each other.

6. Local Authority animation

This animates the log return on a yearly basis and provides a clear visualisation of how much it changes.

7. Local Authority time series decomposition heat maps

Eight different decomposition approaches were investigated. Although the respective trends of median price obtained are always smooth, this is at the cost of placing too much information into the random/remainder/irregular/residual categories, especially during the 2008 crash.

X_13ARIMA_SEATS was used on a quarterly basis for Figure SM.19 and Figure SM.20.

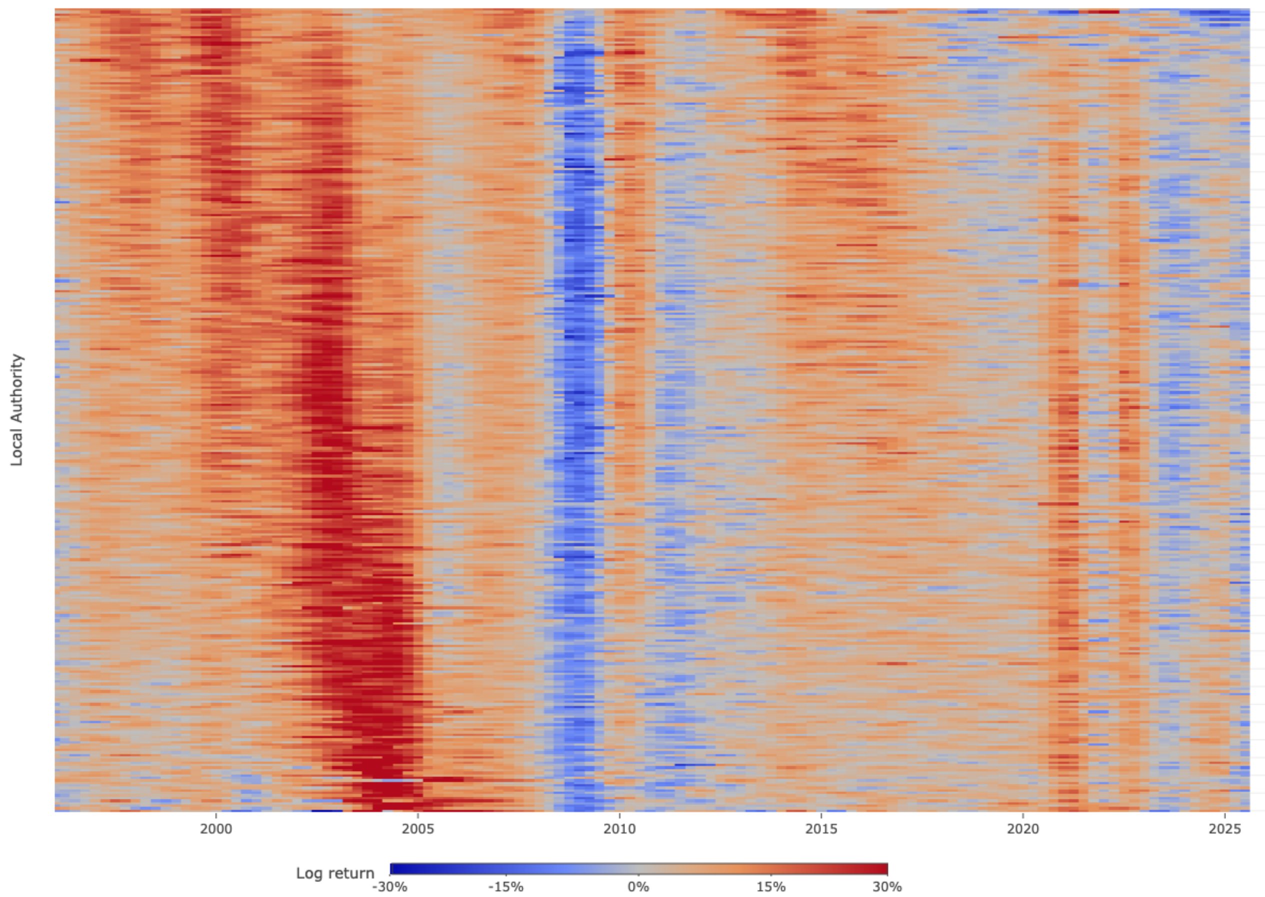

Figure SM.19: Annual log return of the SEATS trend decomposition of the quarterly median, by Local Authority; highest median price Local Authority at the top, lowest median price Local Authority at the bottom.

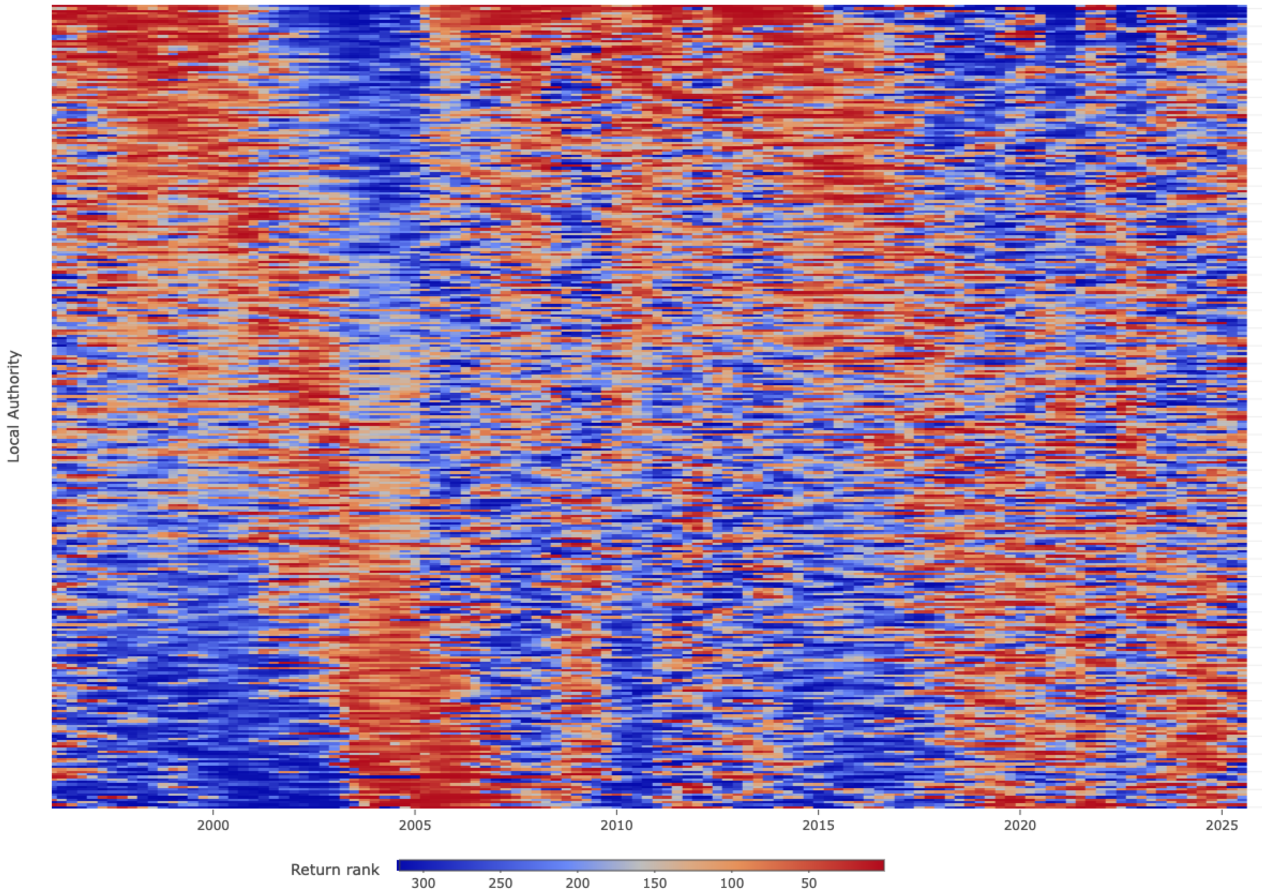

Figure SM.20: Rank of the annual return of the SEATS trend decomposition of the quarterly median, by Local Authority; highest median price Local Authority at the top, lowest median price Local Authority at the bottom.

8. Quantile analysis & 9. Quantile animation

These are the equivalent of 5. Local Authority analysis, and 9. Quantile animation.

In Figure SM.21, the mean and standard deviation of the log returns are shown. The triangle-shaped kinks are due to the old Stamp Duty regime. The standard deviation of the tails is significant and leads to four clusters in the PCA K-means clustering plot in Figure SM.22.

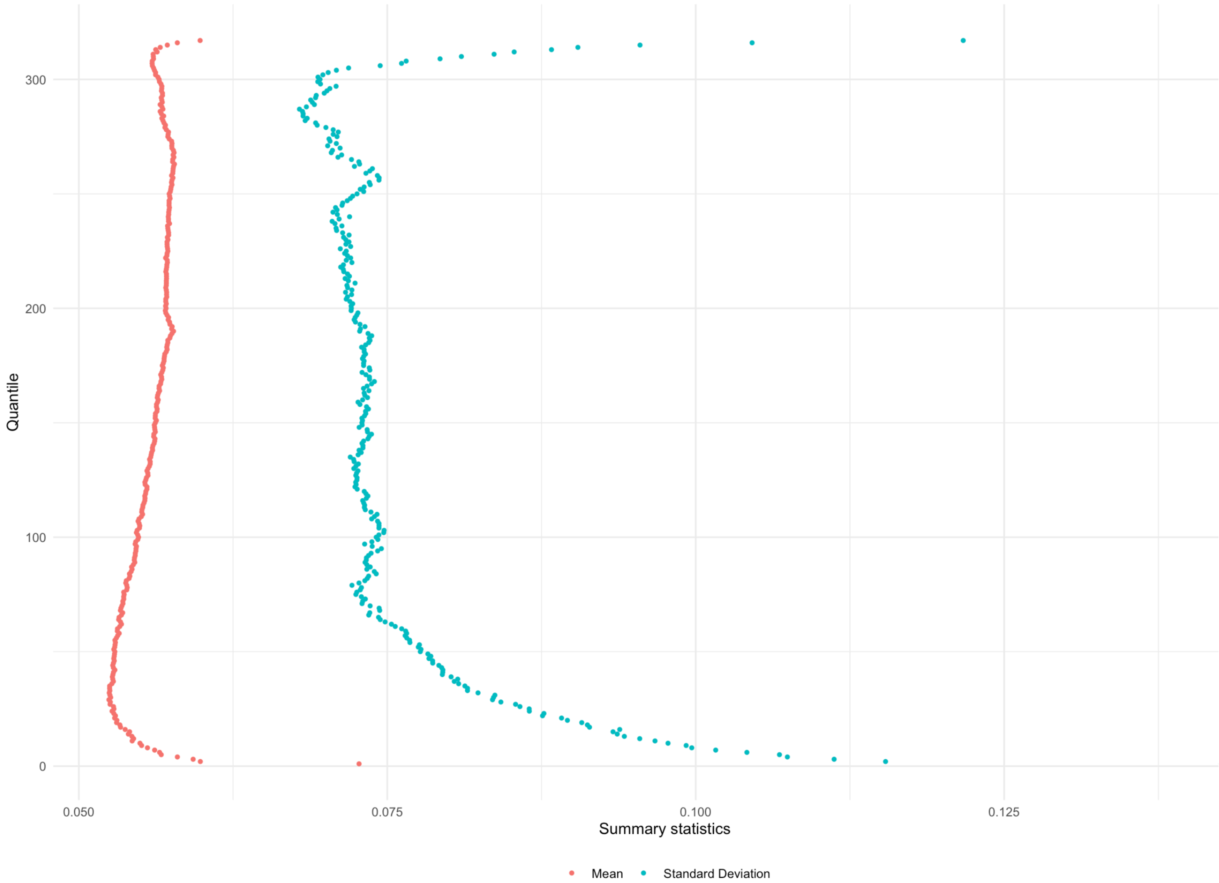

Figure SM.21: Mean and standard deviation of the log return of the median price, by quantile; highest quantile at the top, lowest quantile at the bottom.

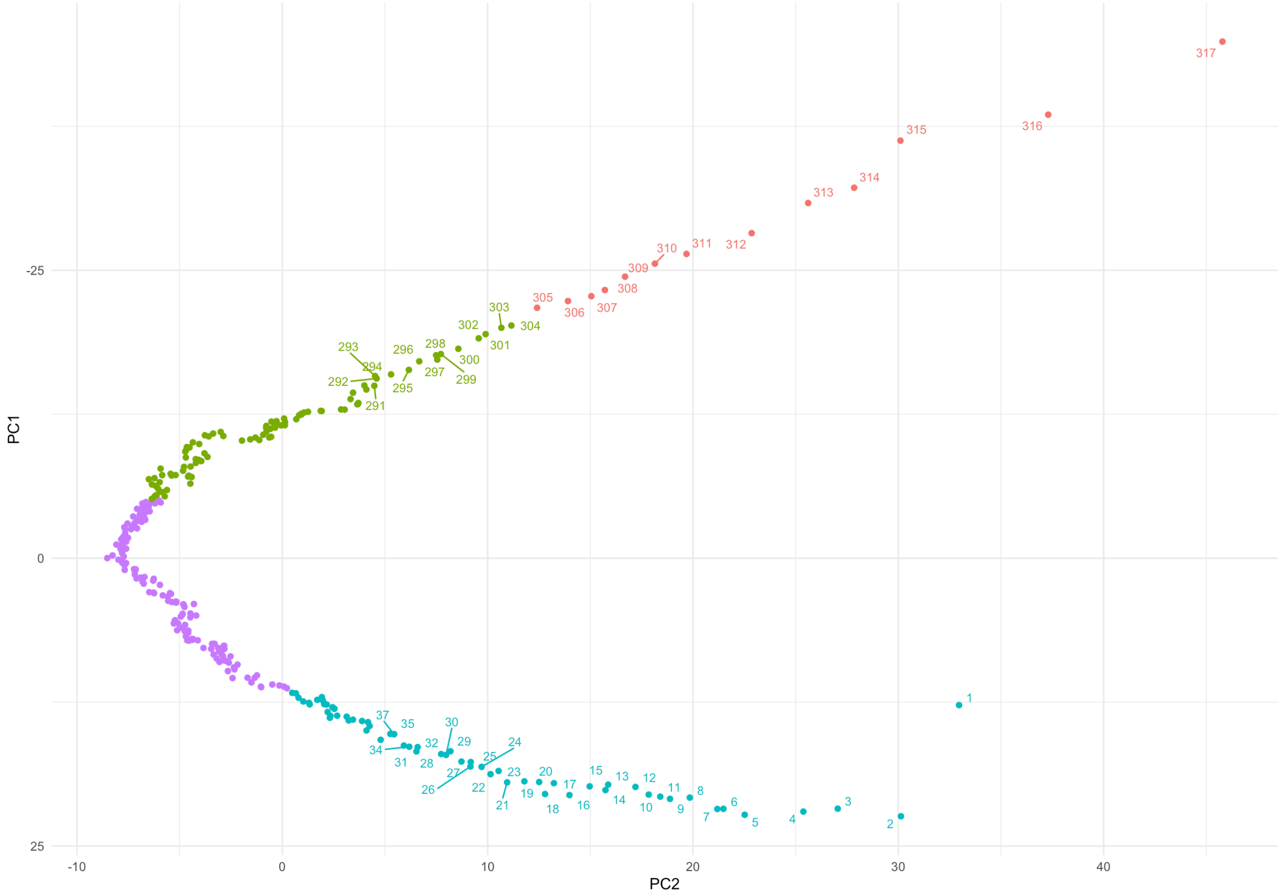

Figure SM.22: Principal Component Analysis K-Means Clustering, by quantile; highest quantile at the top, lowest quantile at the bottom.

10. Credit vs median prices

Are house prices a function of credit[38], or is credit a function of house prices? If you believe that rising house prices are a good thing, then the function is symbiotic when prices are rising, unfavourable when prices are falling: A 'Folie à deux'. Or more formally, a land-credit cycle[39].

The only mortgage data available is for 'Value of residential mortgage loans outstanding in the UK, split by sector postcode'[40]. This data includes house purchases, equity release, remortgaging, and repayments. It is not a like-for-like comparison with sales transactions.

Instead, the following pairs of data from the Bank of England[41] have been used:

- LPMB3C2 Monthly values of total sterling approvals for house purchase to individuals (in sterling millions) not seasonally adjusted
- LPMVTVU Monthly number of total sterling approvals for house purchase to individuals not seasonally adjusted
- LPMB3XF Monthly value of monetary financial institutions' sterling gross approvals for house purchase to individuals (in sterling millions) not seasonally adjusted
- LPMB3ZF Monthly number of monetary financial institutions' sterling gross approvals for house purchase to individuals not seasonally adjusted
- LPMBV86 Monthly value of total sterling gross approvals for house purchase to individuals (in sterling millions) not seasonally adjusted
- LPMBV87 Monthly number of total sterling gross approvals for house purchase to individuals not seasonally adjusted

Added to each value series is the Help to Buy data[42].

Figure SM.23 shows median house prices align with the amount of credit per approval until the pandemic. At different times, there will be variance in the equity mix and number of cash buyers.

Figure SM.23: Moving average median price and moving average total credit per approval

The LPMVTVU dataset has an incorrect number of approvals. In Figure SM.24, the number is actually higher than the number of transactions until 2007-8. The number of approvals should be lower than the number of transactions due to cash buyers. This implies that the LPMB3C2 / LPMVTVU line in Figure SM.23 above should be closer to the median price. This would represent more generous lending standards with lower deposits.

Note that there is a lag of Land Registry registrations for the last few months of the dataset.

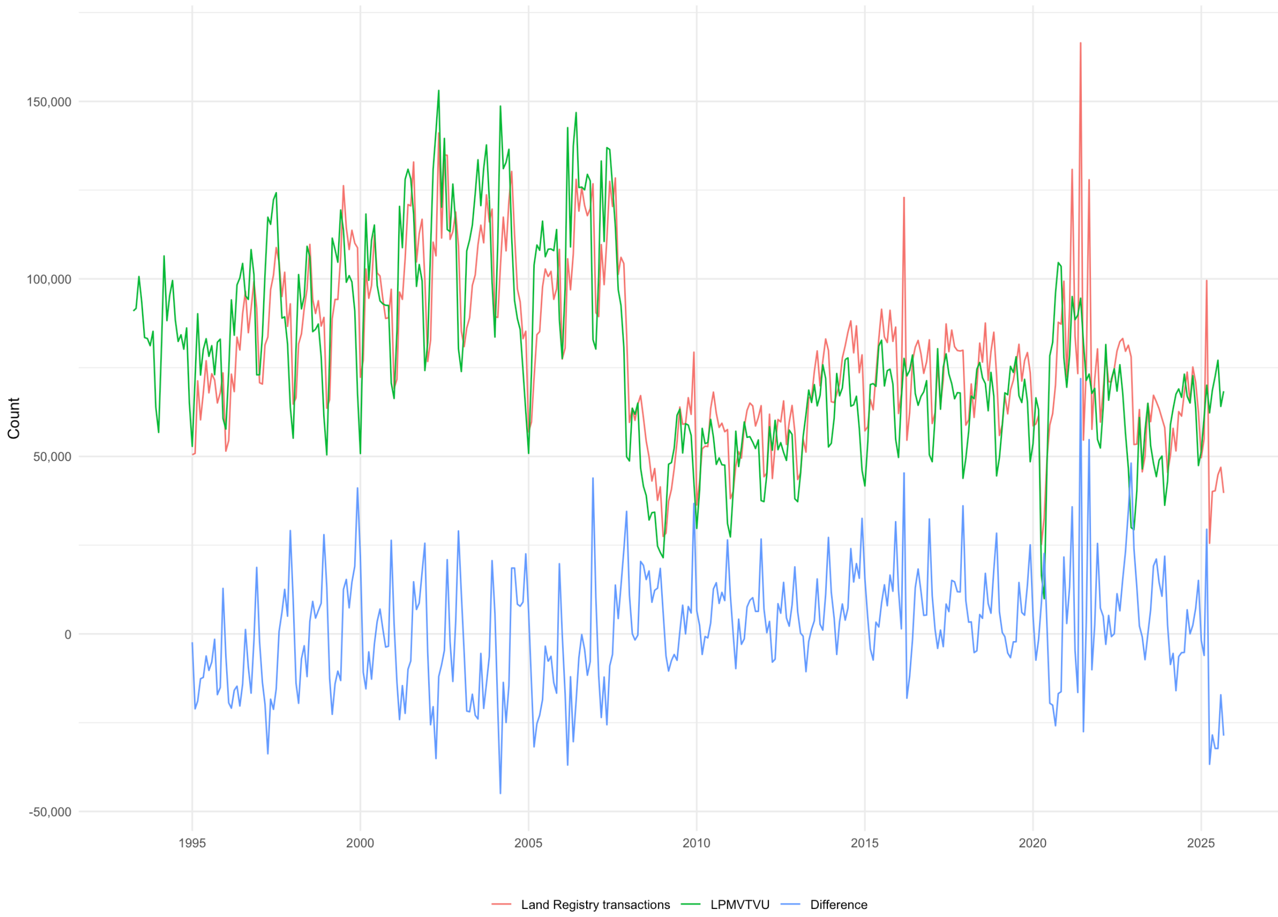

Figure SM.24: Land Registry transactions versus number of credit approvals

In Figures SM.25 and SM.26, two alternate presentations of the median price versus total credit are shown.

In Figure SM.27, the delta of both median house prices and credit is plotted.

Finally, in Figure SM.28, total credit looks to be following a sigmoid curve over the two cycles, although there is no way to know any carrying capacity of the system a priori; furthermore, we are interested in the travelling wave pattern, not the absolute level of prices.

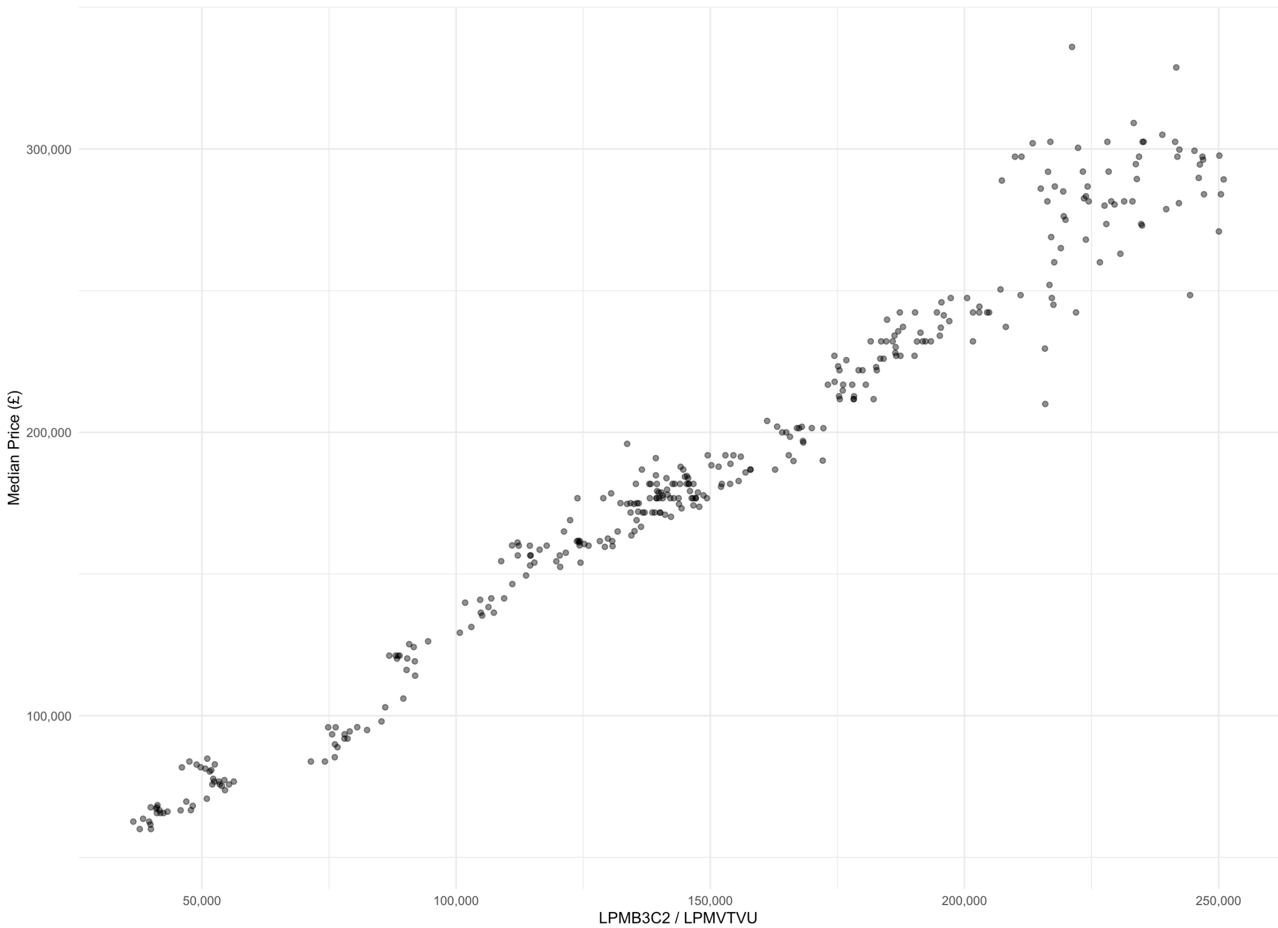

Figure SM.25: Median price versus total credit per approval

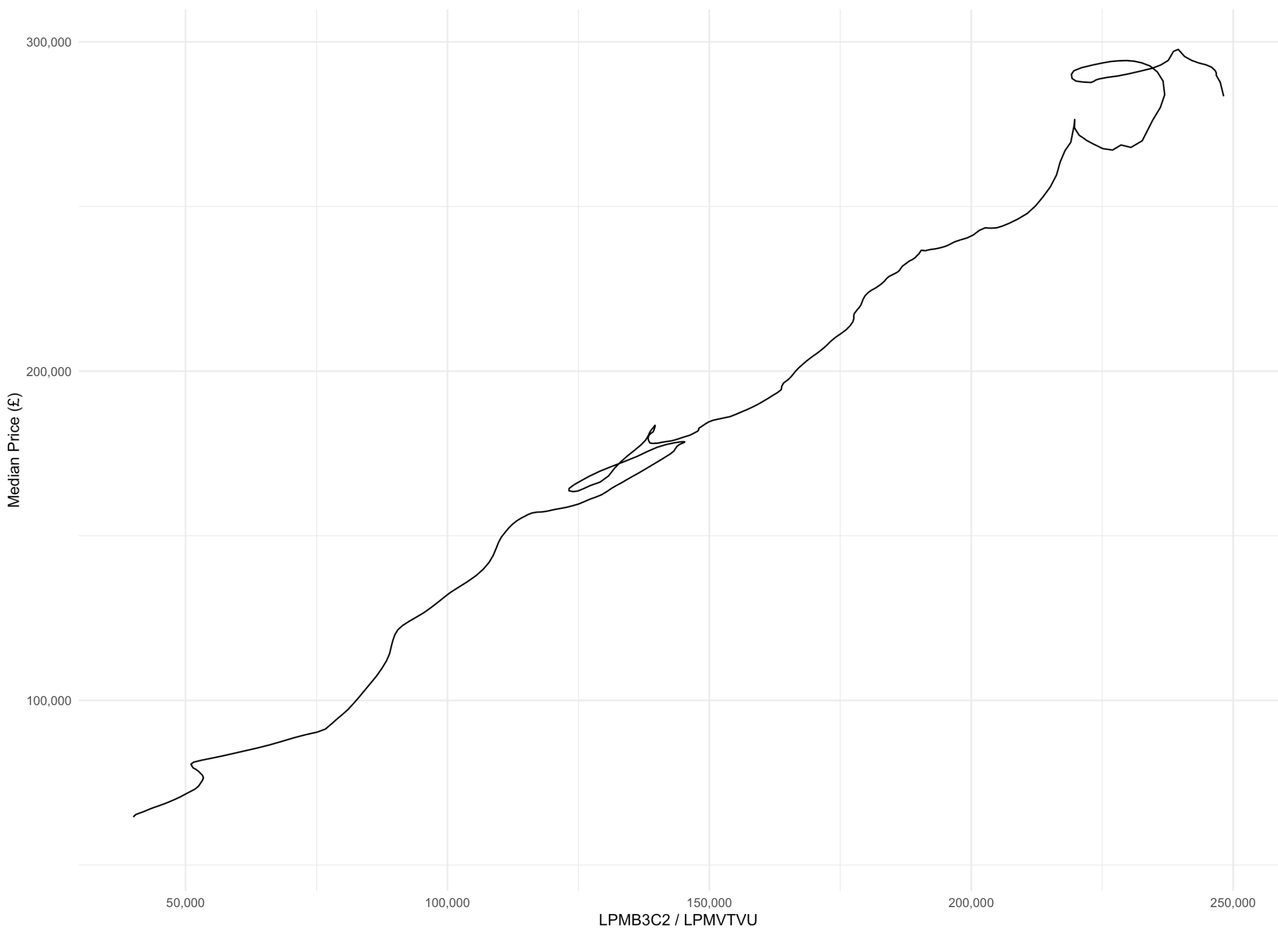

Figure SM.26: Moving average median price versus moving average total credit per approval

Figure SM.27: Median price yearly delta versus credit per transaction yearly delta

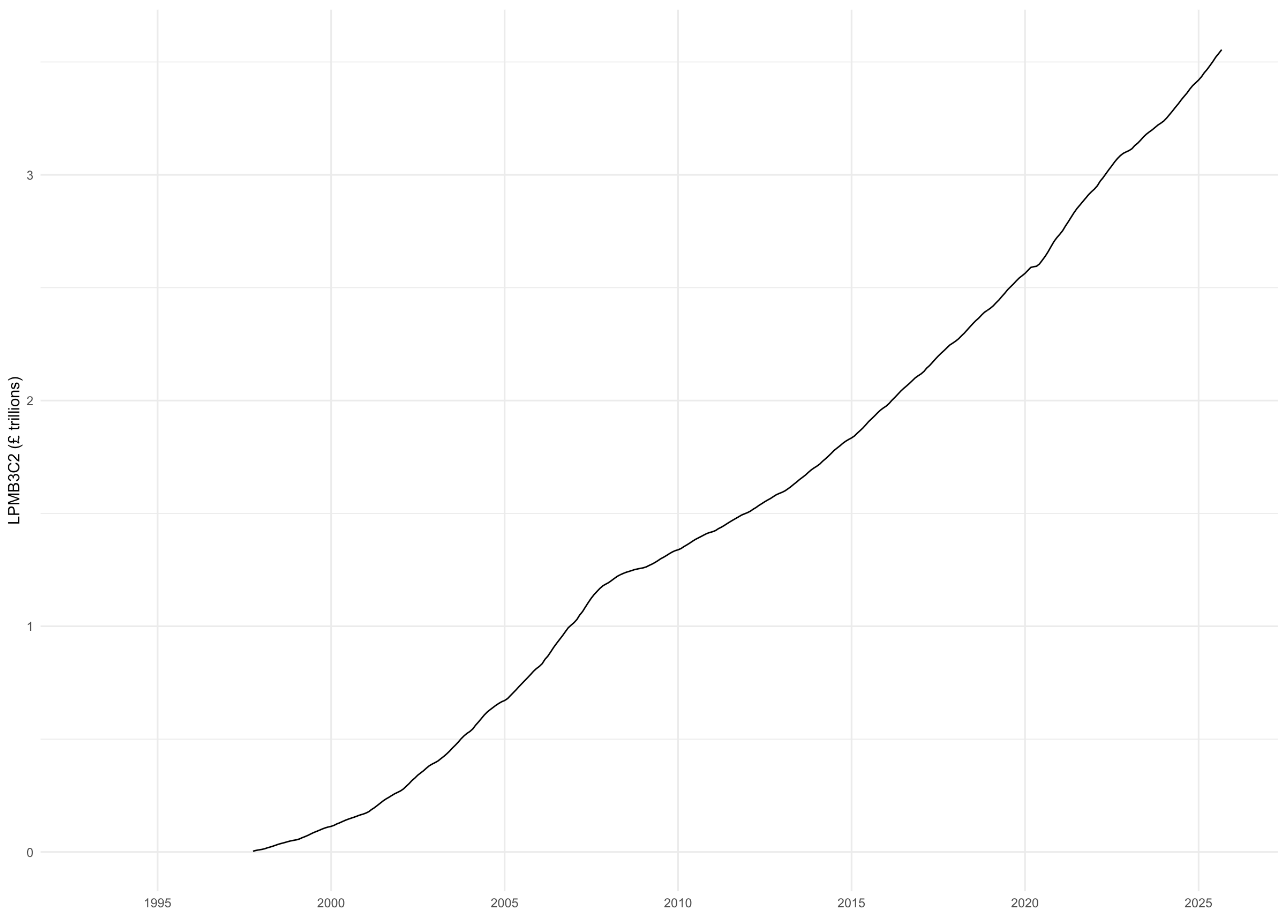

Figure SM.28: LPMB3C2 cumulative credit

11. Repeat sales

We compare repeat sales prices to the first sales price plus inflation from the Consumer Prices Index[43]. Ideally, we would use build costs per region from BCIS[44] but this data is not publicly available.

Sales transactions where there is more than one record for the same Postcode, PAON, SAON, and Street values have been pre-filtered in the script '1. Import and tidy'. As we are only interested in repeat sales at the margin, we do not have to worry about Stamp Duty as these sales are below the lowest threshold and not subject to the tax.

The functions find all pairs of repeat sales, or all triplets of repeat sales; that is, where the same Postcode, PAON, SAON, and Street is available in the two or three years, as appropriate. For simplicity, only the first sale in any given year is included.

The functions interpolate the CPI to the transaction dates and adjust the first, or first and second sales prices to the second, or second and third year.

Results are plotted for the pairs or triplets of sales where the second or third year is 2009. The x-axis is always the first year sales price CPI-adjusted to 2009, and the y-axis is the actual sales price in 2009. A dotted line provides a guide to where the sales price adjusted to 2009 equals the actual 2009 sales price, that is, when the house price has increased by no more than the CPI.

A subset of the returned repeat sales results have been displayed. The reason for this is that we want to avoid plotting results where the house needs renovation or has changed its quantile significantly. For example, if the quantile in 1995 was high but in 2009 it was low, we could suppose that it needs renovation; on the contrary, if the quantile in 1995 was low but in 2009 it was high, it may have been renovated.

As such, we plot repeat sales where the first quantile is from 1 to 5, the last sale is within 4 quantiles of that, and the percentage change in price is within 5%. For the triplet plot, we ensure that the second year's sales price is higher than the first and third sales prices. This can be adjusted as required, but ultimately, we are searching for the repeat transactions that meet these criteria at the lowest price, which we suggest is a property where the land value is zero and the material costs form the basis of the property value in the 2009 trough.

There are three houses at the bottom left of the Figure SM.29, and either of these could serve as a good example, but their value in the time between their first sale and 2009 is not known in this pari analysis.

At the bottom of Figure SM.30, there are two labels for a single point. The code has actually found repeat sales in four different years for the same property: in 2002 (£15,750), 2004 (£36,000), 2006 (£35,500), and 2009 (£18,000). Here is an example of a house at the margin that experienced the boom only to be worth its 2002 cost plus inflation in the trough. Alternate examples at slightly higher price points can also be found.

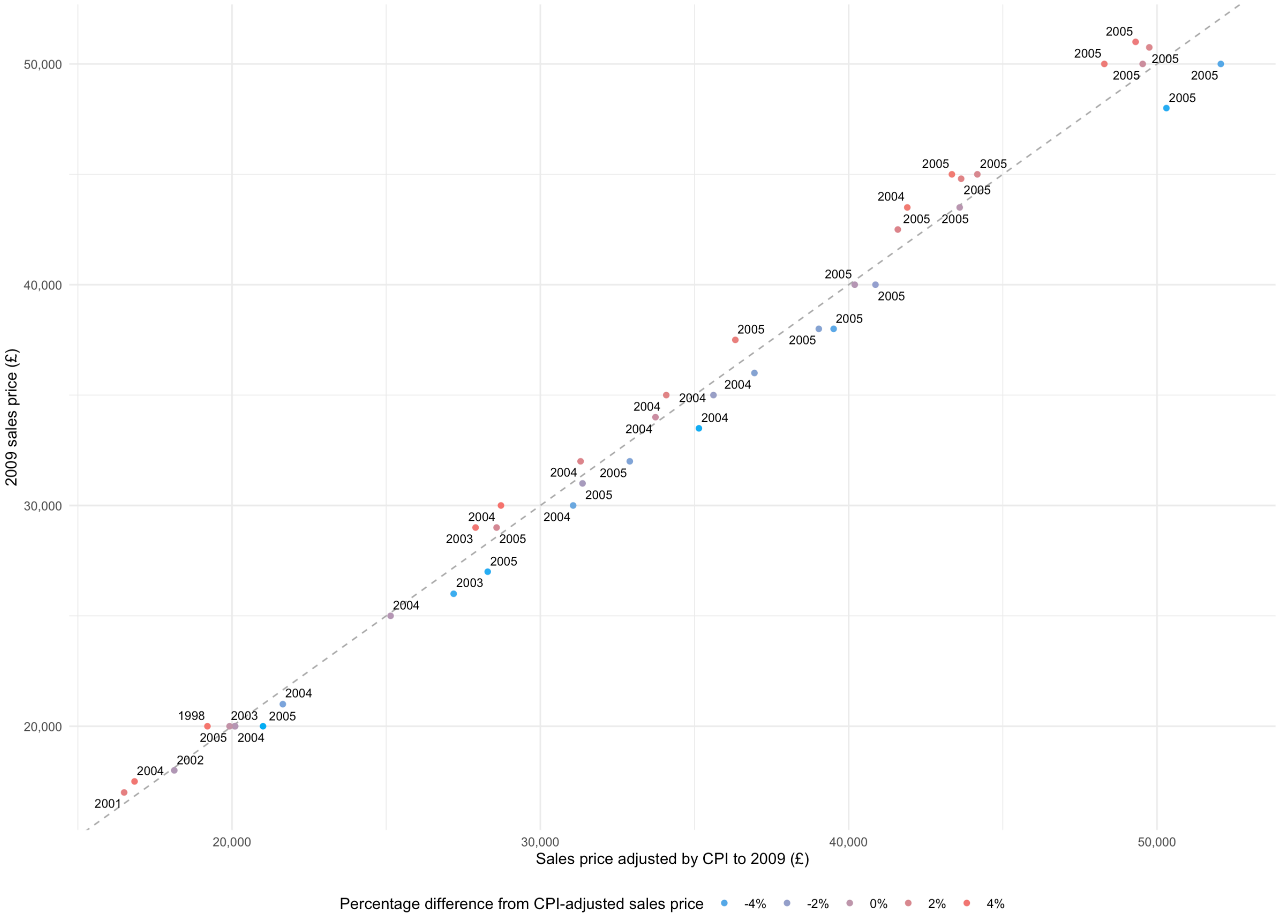

Figure SM.29: Repeat sales - sales prior to 2006 CPI-adjusted to 2009 versus 2009 sales prices

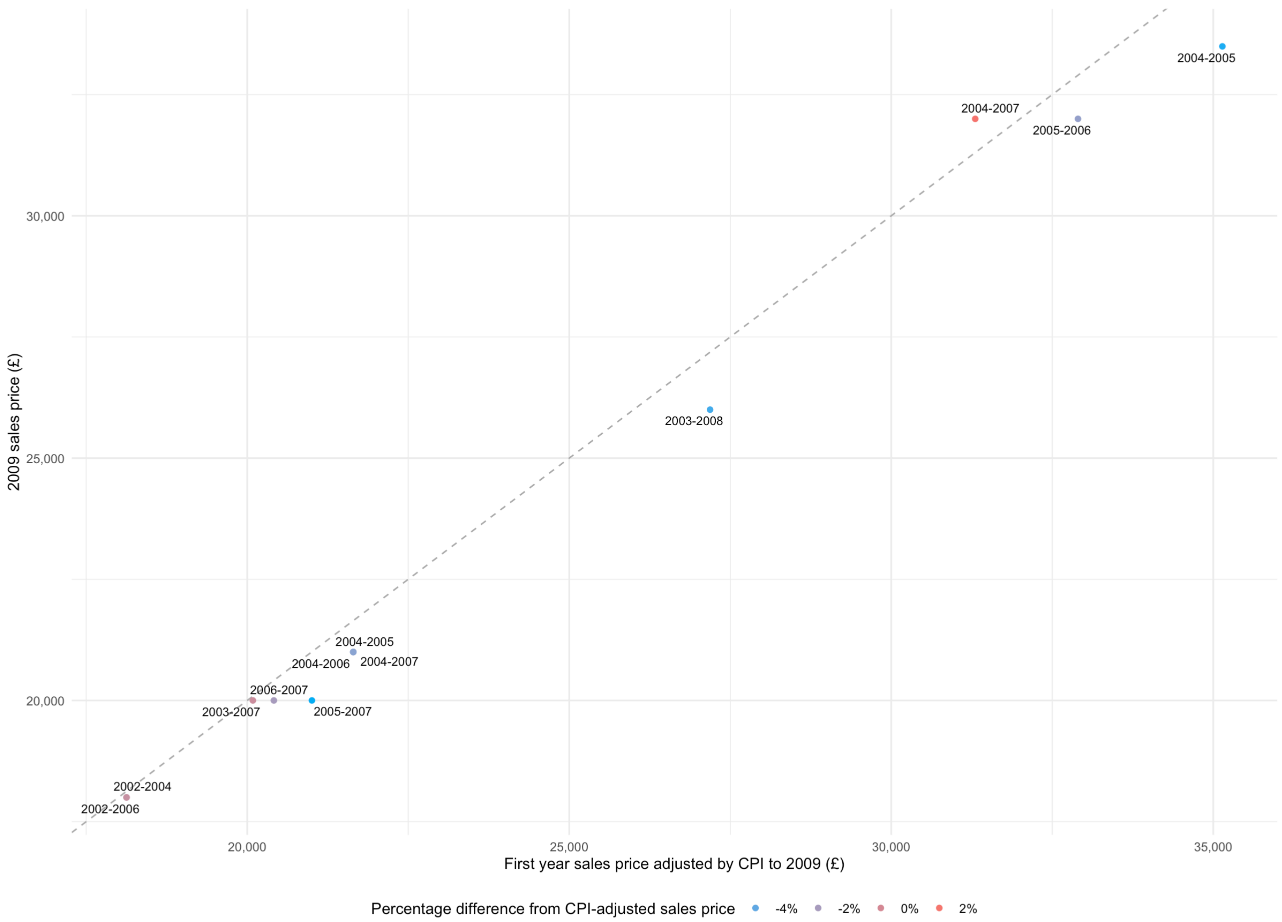

Figure SM.30: Repeat sales - multiple sales with first year CPI-adjusted to 2009 versus 2009 sales prices

12. Local Authority price-to-earnings ratio

In Figure SM.31, the ONS House price to workplace-based earnings ratio[45] has been plotted against Local Authority.

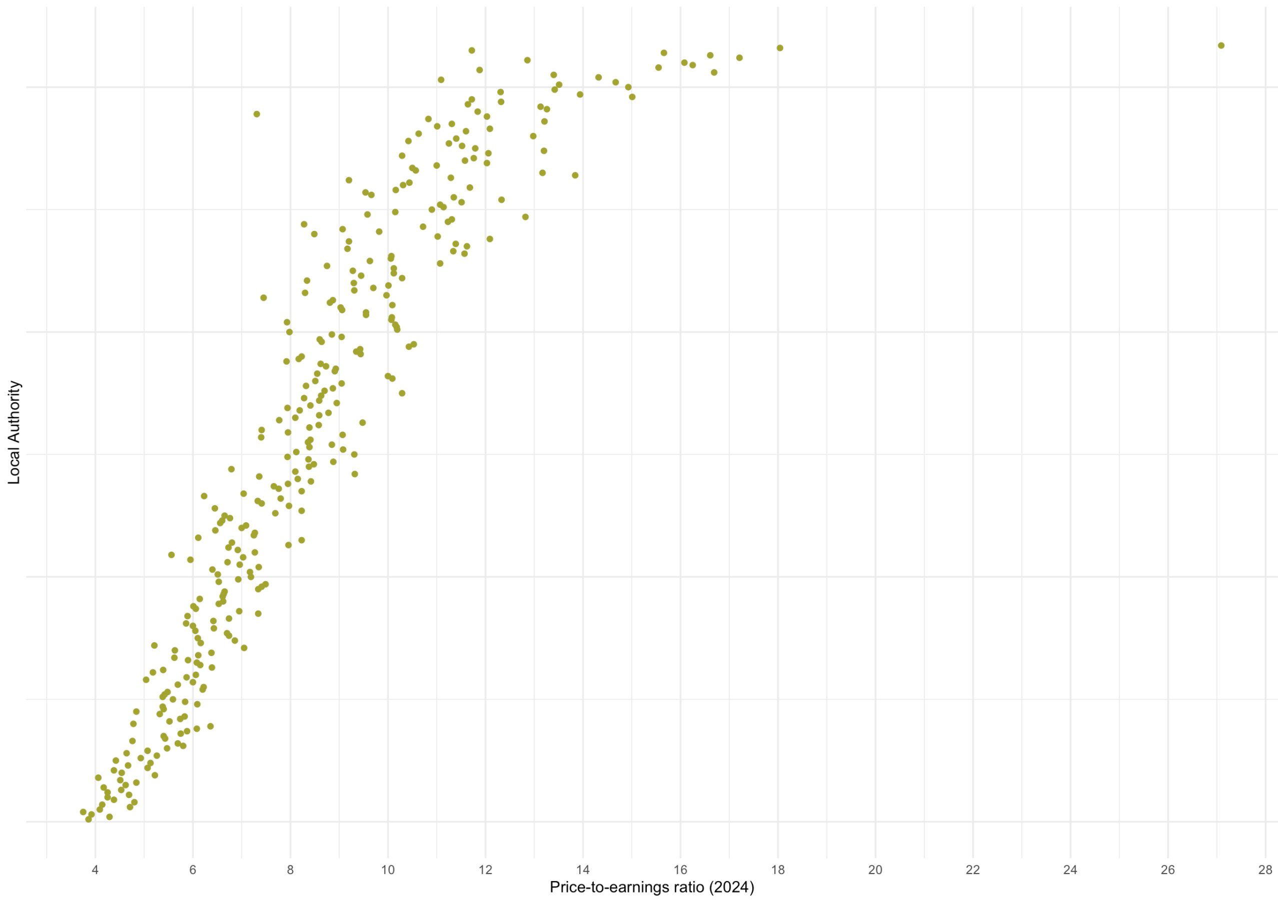

Figure SM.31: Price-to-earnings ratio (2024), by Local Authority; highest median price Local Authority at the top, lowest median price Local Authority at the bottom.

C. Use of Large Language Models

Code

All generated code was based on the author's ideas, with author iteration thereafter.

OpenAI ChatGPT

2. Stamp Duty

- JSON approach to thresholds and tax rates
- Stamp Duty calculation functions

3. Yearly distributions

- Correct colour assignment in the Cullen and Frey plot

5. Local Authority analysis & 8. Quantile analysis

- Principal Component Analysis, K-Means Clustering
- plotly legend: ordering and 2x2 layout

For the price per square metre dataset:

3. PPD_EPC_linkage fast

- beg2char() and char2end() equivalents using stringi

4. Create final PPD

- data.table syntax

Anthropic Claude

5. Local Authority analysis & 8. Quantile analysis

- Use of the igraph package, specifically the graph_from_data_frame() and as_adjacency_matrix() functions, to plot the spatiotemporal rank analysis

11. Repeat sales

- Year combination and analysis functions

For the price per square metre dataset:

1. Combine EPC certificates

- Iterative functions with error handling

Narrative

Apple Intelligence

In-line proofreading in Pages: tidying up commas, colons, and stray words.

OpenAI ChatGPT & Google Gemini

Suggestions for tightening the narrative flow, mainly removing sections. Suggestions for the Abstract and Conclusion to improve the academic tone.

D. About the Author

I graduated with an MBA into the 2008 great financial crisis. This experience led me to want to understand the root cause of the housing market crash.

In subsequent years, I took a medium-sized London property project through planning; initially 14 units, later 27 units. This was a perfect education to the role of land in the economy.

https://uk.linkedin.com/in/chris-beech-0

E. Acknowledgements

Had I not read Martin's Wolf's Financial Times column[46] advocating a land value tax and referencing the work of Fred Harrison on the 18-year land cycle[47], this work may never have existed. I have learnt much from both over the years, thank you.

Thank you to Neal Hudson for allowing the inclusion of his plot; a simple visual has inspired so much.

Thank you to Bin Chi for a momentous effort linking Land Registry and EPC records to create a price per square metre dataset.

Thank you to the authors of all the R software used in the creation of this work.

Thank you to those who listened to these ideas over many years and those who reviewed this work for narrative clarity. All mistakes are my own.

F. Funding statement

This work was done in my spare time with no grant support.

G. Competing interests

I have no competing interests.

R Library citations

Barrett T, Dowle M, Srinivasan A, Gorecki J, Chirico M, Hocking T, Schwendinger B, Krylov I (2025). *data.table: Extension of `data.frame`* doi:10.32614/CRAN.package.data.table https://doi.org/10.32614/CRAN.package.data.table, R package version 1.17.2, https://CRAN.R-project.org/package=data.table

Csárdi G, Nepusz T, Traag V, Horvát Sz, Zanini F, Noom D, Müller K (2025). *igraph: Network Analysis and Visualization in R*. doi:10.5281/zenodo.7682609 https://doi.org/10.5281/zenodo.7682609, R package version 2.1.4, https://CRAN.R-project.org/package=igraph

Delignette-Muller, M L, Dutang, C (2015). *fitdistrplus: An R Package for Fitting Distributions*. Journal of Statistical Software, 64(4), 1-34. DOI 10.18637/jss.v064.i04.

Firke S (2024). *janitor: Simple Tools for Examining and Cleaning Dirty Data*. doi:10.32614/CRAN.package.janitor https://doi.org/10.32614/CRAN.package.janitor, R package version 2.2.1, https://CRAN.R-project.org/package=janitor

Gagolewski M., *stringi: Fast and portable character string processing in R*, Journal of Statistical Software 103(2), 2022, 1–59, doi:10.18637/jss.v103.i02.

Hyndman R (2024). *fpp3: Data for "Forecasting: Principles and Practice" (3rd Edition)* doi:10.32614/CRAN.package.fpp3 https://doi.org/10.32614/CRAN.package.fpp3, R package version 1.0.1, https://CRAN.R-project.org/package=fpp3

Mühleisen H, Raasveldt M (2025). *duckdb: DBI Package for the DuckDB Database Management System*. doi:10.32614/CRAN.package.duckdb https://doi.org/10.32614/CRAN.package.duckdb, R package version 1.4.1, https://CRAN.R-project.org/package=duckdb

O'Hara-Wild M (2020). *fable.prophet: Prophet Modelling Interface for 'fable'*. doi:10.32614/CRAN.package.fable.prophet https://doi.org/10.32614/CRAN.package.fable.prophet, R package version 0.1.0, https://CRAN.R-project.org/package=fable.prophet

Ooms J (2014). *The jsonlite Package: A Practical and Consistent Mapping Between JSON Data and R Objects* arXiv:1403.2805 [stat.CO] https://arxiv.org/abs/1403.2805

Pebesma, E., 2018. *Simple Features for R: Standardized Support for Spatial Vector Data*. The R Journal 10 (1), 439-446, https://doi.org/10.32614/RJ-2018-009

R Core Team (2025). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

R Special Interest Group on Databases (R-SIG-DB), Wickham H, Müller K (2024). *DBI: R Database Interface*. doi:10.32614/CRAN.package.DBI https://doi.org/10.32614/CRAN.package.DBI, R package version 1.2.3, https://CRAN.R-project.org/package=DBI

Sax C, Eddelbuettel D (2018). *Seasonal Adjustment by X-13ARIMA-SEATS in R* Journal of Statistical Software, *87*(11), 1-17. doi:10.18637/jss.v087.i11 https://doi.org/10.18637/jss.v087.i11

Sievert C. *Interactive Web-Based Data Visualization with R, plotly, and shiny*. Chapman and Hall/CRC Florida, 2020.

Slowikowski K (2024). *ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2'* doi:10.32614/CRAN.package.ggrepel https://doi.org/10.32614/CRAN.package.ggrepel, R package version 0.9.6, https://CRAN.R-project.org/package=ggrepel

Vaughan D (2024). *slider: Sliding Window Functions* doi:10.32614/CRAN.package.slider https://doi.org/10.32614/CRAN.package.slider, R package version 0.3.2, https://CRAN.R-project.org/package=slider

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019). *Welcome to the tidyverse* Journal of Open Source Software, *4*(43), 1686. doi:10.21105/joss.01686 https://doi.org/10.21105/joss.01686

Wickham H, Bryan J (2025). *readxl: Read Excel Files* doi:10.32614/CRAN.package.readxl https://doi.org/10.32614/CRAN.package.readxl, R package version 1.4.5, https://CRAN.R-project.org/package=readxl

Wickham H, Pedersen T, Seidel D (2025). *scales: Scale Functions for Visualization* doi:10.32614/CRAN.package.scales https://doi.org/10.32614/CRAN.package.scales, R package version 1.4.0, https://CRAN.R-project.org/package=scales

All code was formatted using Air:
https://www.tidyverse.org/blog/2025/02/air/

References

[1] Beech, C.
https://github.com/chris-beech-housing

All code and this paper is available on GitHub. This version: Saturday, 6 December 2025.

[2] Lackman, C.L., 1976. The classical base of modem rent theory. American Journal of Economics and Sociology, 35(3), pp.287-300.

[3] Ricardo, D., 1821. On the Principles of Political Economy and Taxation. London: J. Murray.

[4] Knoll, K., Schularick, M. and Steger, T., 2017. No price like home: Global house prices, 1870–2012. American Economic Review, 107(2), pp.331-353.

[5] Ryan-Collins, J., Lloyd, T. and Macfarlane, L., 2017. Rethinking the economics of land and housing. Bloomsbury Publishing.

[6] Meen, G., 1999. Regional house prices and the ripple effect: a new interpretation. Housing Studies, 14(6), pp.733-753.

Historically, ripple effect papers have used house price indices from Halifax and/or Nationwide for the regions of the UK.

[7] Cook, S. and Watson, D., 2016. A new perspective on the ripple effect in the UK housing market: Co-movement, cyclical subsamples and alternative indices. Urban Studies, 53(14), pp.3048-3062.

[8] Chi, B., Dennett, A., Oléron-Evans, T. and Morphet, R., 2022. Delineating the Spatio-Temporal Pattern of House Price Variation by Local Authority in England: 2009 to 2016. Geographical Analysis, 54(2), pp.219-238.

[9] Hoyt, H., 2000. One hundred years of land values in Chicago: The relationship of the growth of Chicago to the rise of its land values, 1830-1933. Beard Books.

[10] Foldvary, F.E., 1997. The Business Cycle: A Georgist-Austrian Synthesis. American Journal of Economics and Sociology, 56(4), pp.521-524.

[11] Harrison, F., 2010. Boom Bust: house prices, banking and the depression of 2010. Shepheard-Walwyn

[12] HM Land Registry Price Paid Data
https://www.gov.uk/guidance/about-the-price-paid-data

[13] ONS Postcode Directory (August 2025) for the UK
https://geoportal.statistics.gov.uk/datasets/3be72478d8454b59bb86ba97b4ee325b/about

Local Authorities have been consolidated over time, so it is important to use the latest mapping.

[14] Local Authority Districts (April 2025) Names and Codes in the UK
https://www.data.gov.uk/dataset/b2c91962-58e7-40f1-ad56-7aa2473a93fd/local-authority-districts-april-2025-names-and-codes-in-the-uk-v21

[15] Beech, C.
https://github.com/chris-beech-housing/Land-Registry-PPD-EPC-price-per-square-metre

This code refactor is 18 times faster than Chi's original version.

[16] Chi, B., Dennett, A., Oléron-Evans, T. and Morphet, R., 2021. A new attribute-linked residential property price dataset for England and Wales, 2011–2019.

Chi has matched PPD to EPC certificates to obtain a size for each property; match percentages improve from 63.9% in 1995 to 97.4% in 2025; this is equivalent to a random sample so does not affect the results.

[17] Leamer, E.E., 2007. Housing *is* the business cycle.

Leamer has a passing mention of quantiles, but our approach was chosen independently when thinking about the wide range of house prices in each Local Authority.

[18] This heat map is no longer available online but is included in the Supplementary Material.

[19] Ibid. 11

[20] Bank of England, How much of the recent house price growth can be explained by the 'race for space'? https://www.bankofengland.co.uk/bank-overground/2021/how-much-of-the-recent-house-price-growth-can-be-explained-by-the-race-for-space

[21] Scheffer, M., Bascompte, J., Brock, W.A., Brovkin, V., Carpenter, S.R., Dakos, V., Held, H., Van Nes, E.H., Rietkerk, M. and Sugihara, G., 2009. Early-warning signals for critical transitions. Nature, 461(7260), pp.53-59.

This paper discusses 'Spatial patterns as early-warning signals.'

[22] Krainer, J., LeRoy, S.F., Munpyung, O., 2009. Mortgage Default and Mortgage Valuation. Federal Reserve Bank of San Francisco, Working Paper 2009-20 http://www.frbsf.org/publications/economics/papers/2009/wp09-20bk.pdf

[23] Gardiner, C.W., 2021. Elements of stochastic methods. Melville, NY, USA: AIP Publishing.

[24] LLMs can provide a very good Fokker-Planck mathematical formalisation of this model, but we prefer to focus on a Ricardian explanation; the intuition of thinking about land value is more important.

[25] Wolf, M., 2010, Why we must halt the land cycle https://www.ft.com/content/8f06df9e-8ac1-11df-8e17-00144feab49a

[26] Gaffney, M., 1982. Two centuries of economic thought on taxation of land rents. Land Value Taxation in Thought and Practice, pp.151-196.

[27] Kumhof, M., Tideman, T.N., Hudson, M. and Goodhart, C., 2021. Post-corona balanced-budget super-stimulus: The case for shifting taxes onto land. CEPR Discussion Papers.

[28] Patel, K., Paxson, D., Sing, T.F., 2005. RICS Research Paper Series, Volume 5, Number 1, Review of the Practical Uses of Real Property Options

We view housing as three different markets, where price is best modelled as follows:
- The development land market is an options market
- The rental market using supply and demand
- The sales market using supply and demand of credit; if credit grows by more than the supply of houses then it is credit that drives house prices, Ibid. 5

[29] Ibid. 17

[30] LSE (8th September 2022), 'Her Majesty Elizabeth II, A tribute', London School of Economics https://www.lse.ac.uk/alumni-friends-and-partners/Her-Majesty-Elizabeth-II

Queen Elizabeth II famously asked, "Why did no one see it coming?"

[31] Privacy Impact Assessment review: Price Paid Data
https://www.gov.uk/government/publications/privacy-impact-assessment-review-price-paid-data

The UK assessment is that, “Price Paid Data is not personal information but property related information.” We can only assume that other governments disagree and, as a result, do not publish their land registry data.

We note that the analysis by quantile only requires a date and a price, ideally a price per square metre.

[32] RICS, Valuation of development property, 1st edition, October 2019
https://www.rics.org/content/dam/ricsglobal/documents/to-be-sorted/Valuation%20of%20development%20property_ready%20for%20approvals.pdf

[33] History of Stamp Duty Taxes
https://www.stampdutyrates.co.uk/historic-rates.html

[34] Stamp Duty Land Tax rates from 1 December 2003 to 31 March 2005
https://www.gov.uk/government/publications/rates-and-allowances-stamp-duty-land-tax

[35] Land Transaction Tax rates and bands
https://www.gov.wales/land-transaction-tax-rates-and-bands

[36] Local Authority Districts (May 2025) Boundaries UK BFE
https://geoportal.statistics.gov.uk/datasets/ons::local-authority-districts-may-2025-boundaries-uk-bfe-v2/about

[37] Dorling, D. 2010. Persistent North-South divides, Chapter 2 in N.M. Coe and A. Jones (Eds) The Economic Geography of the UK, London: Sage, pp12-28.

[38] McLeay, M., Radia, A. and Thomas, R., 2014. Money creation in the modern economy. Bank of England, Quarterly Bulletin 2014 Q1
https://www.bankofengland.co.uk/quarterly-bulletin/2014/q1/money-creation-in-the-modern-economy

[39] Ibid. 5

[40] UK Finance, Mortgage lending within UK postcodes
https://www.ukfinance.org.uk/data-and-research/data/mortgage-lending-within-uk-postcodes

[41] Bank of England
https://www.bankofengland.co.uk/boeapps/database/

[42] Help to Buy Tables
https://assets.publishing.service.gov.uk/media/64abb06ca32f130013f06907/Help_to_Buy_Tables.ods

This data has been simplified for import.

[43] Consumer price inflation time series (MM23), CPI INDEX 00: ALL ITEMS 2015=100
https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/d7bt/mm23

[44] BCIS Construction Indices
https://bcis.co.uk/insight/construction-indices-cost-and-price-indices/

[45] ONS House price to workplace-based earnings ratio
https://www.ons.gov.uk/peoplepopulationandcommunity/housing/datasets/ratioofhousepricetoworkplacebasedearningslowerquartileandmedian

[46] Ibid. 25

[47] Ibid. 11