REGIONAL HOUSING WEALTH INEQUALITY

The Big Data Perspective

Timofey Sugaipov

Abstract

The aim of this paper is to understand how regional housing wealth inequality has evolved in England over the past 50 years. The paper attempts to measure regional inequality on the national level, as well as focusing on inequality within regions during the later periods for which more data is available. This paper is intended as a contribution to the literature on regional housing wealth inequality. While housing wealth inequality in England has been explored before, few studies have investigated variations within regions. Furthermore, our dataset, price paid data from the HM Land Registry, is quite unique since it represents the population of property transactions during the period of 1995-2018 rather than the usual sample from surveys. This dataset has been used before in earlier studies of housing wealth inequality however the amount of data available during those studies was much smaller.

Introduction

The aim of this paper is to understand how regional housing wealth inequality has evolved in England over the past 50 years. The paper attempts to measure regional inequality on the national level, as well as focusing on inequality within regions during the later periods for which more data is available. This paper is intended as a contribution to the literature on regional housing wealth inequality. While

housing wealth inequality in England has been explored before, few studies have investigated variations within regions. Furthermore, our dataset, price paid data from the HM Land Registry, is quite unique since it represents the population of property transactions during the period of 1995-2018 rather than the usual sample from surveys. This dataset has been used before in earlier studies of housing wealth inequality however the amount of data available during those studies was much smaller.

Studies of inequality are quite established in economic literature, stretching back as far as Malthus and Marx. However, studies of regional inequality, inequality within the country, are much more recent. Regional inequality is quite peculiar as it implies that areas with quite similar attributes can experience dramatically different development. The North-South divide in the UK is a perfect example of this, with the North being “left-behind” while the South prospered at her expense. This ideology is held amongst many in the North and breeds discontent with the current political system allowing a rise of radical ideology within the region, as can be seen with the recent Brexit vote. There have been numerous studies about the North-South divide and its causes. However, few studies have looked at housing wealth.

This is strange as wealth can have seen us the most relevant measure of inequality. Housing wealth inequality is the most visible type of inequality, as housing is one of the most tangible types of asset and can be seen by anyone, while other types of assets are much harder to gauge. As a result, rising housing wealth inequality is likely to result in increasing tensions within the country, due to the visible wealth gap becoming more prominent. Therefore, studying housing wealth inequality within regions is relevant, since it can provide us with a measure of discontent within the region, which is usually an area of that is overlooked in studies of regional housing wealth inequality.

While it may seem that housing wealth inequality is only relevant to homeowners, that is far from the truth. Since rising housing wealth inequality could imply the rising house prices which reduce the

availability of affordable housing within the region, which is usually matched by rising rents within the region. This effect is relevant as it reduces the disposable income of renters, which lowers their welfare. London's housing crisis is a good example of this effect in action.[1]

There is an absence of reliable primary sources of regional housing wealth data, therefore we will use the closest proxy, housing prices. Due to the length of our study, it would be fair to assume that variations in market transactions, in the long run, are representative of changes in the overall wealth of the household. This is supported by Lewin and Price who argue that changes in the distribution of housing prices have significant effects on the overall wealth of the household.[2] The paper by Lewin and Pryce is quite valuable for our analysis as they use a similar methodology for calculating housing wealth inequality, although the period they cover is much shorter. In a way, part of this paper is the continuation of their study. There are several other papers that had an influence on this study which will be covered in the next section.

Historiographical Context

One of the most famous publications on wealth inequality is "Capital in the Twenty-First Century" by Thomas Piketty. This book puts wealth inequality into a global perspective and demonstrates how it has evolved over time. While the book itself covers a lot of areas, the topic of housing wealth seems to be underrepresented overall. Nonetheless, Piketty demonstrates that the wealth shares of the top 10% of households increased from 65% to over 70% in the 2010s.[3] This shows that wealth inequality has been increasing over the recent period of history. This statement is further supported by the analysis of the

[1] Travers T., Sims S., Bosetti N., "Housing Inequality in London", Center for London, 2016: 24
[2] Levin E. & Price G., "Measuring Changes in Housing Wealth Inequality", Department for Communities and Local Government, Economics paper 6: Volume 2, 2010
[3] Piketty T., "Capital in the 21st Century", Cambridge: Harvard University Press, 2014

US economy by Saez and Zucman.[4] Although Piketty puts emphasis on the idea that the distribution of wealth is determined more by politics than economic or social factors. While this might be the case for the distribution of wealth, we would argue that the distribution of housing wealth is influenced more by market forces, due to the majority of wealth gains being generated from rising market values of the property. While Piketty's analysis is quite useful in helping us understand the general mechanics of wealth inequality over the period, we require a perspective that deals with wealth inequality in the UK rather than the world.

This is where Atkinson's perspective becomes invaluable. In his book "Unequal Shares", Atkinson argues the inequality in the current distribution of wealth leads to unequal incomes due to the additional income one can receive from wealth.[5] This can lead to a further deterioration of the wealth distribution generating a cycle of rising inequality. He provides several options for redistribution of wealth, to tackle the extreme degree of wealth inequality in Britain.[6] While his suggestions could be effective, they are constrained by politics, which supports Piketty's original argument. Atkinson does show that housing and land wealth account for the largest portion of total personal wealth in Britain in 1966.[7] Even though this measure is quite dated, it still demonstrates the idea that housing wealth accounts for most of the household wealth in Britain. Implying that wealth and housing wealth inequality are strongly linked, meaning that rising housing wealth inequality can be prone to a similar cycle of rising inequality. While housing does have low liquidity, possible changes in rent incomes can still fuel the cycle. Therefore, we will shift focus to literature that covers housing wealth inequality.

There is a surprisingly small amount of literature on housing wealth inequality in England. Notably, there

[4] Saez E. & Zucman G., "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data (No. W20625)", National Bureau of Economic Research, 2014
[5] Atkinson A.B., "Unequal Shares", Allen Lane The Penguin Press, 1972, 251
[6] Ibid
[7] Atkinson A.B., "Unequal Shares", Allen Lane The Penguin Press, 1972, 7

is a sizable body of literature on housing wealth in general, but few investigates changes in housing wealth inequality over time.[8] Although the recent housing crisis has resulted in this topic getting more coverage, this area is still under-explored.

The first study of UK regional wealth inequality was done by Henley.[9] The study utilised data from two surveys: The General Household Survey (GHS) and the British Household Panel Survey (BHPS). Henley specifically looked at movements in the housing wealth distribution and how it was affected by changes in the determinants of household housing wealth. Henley's analysis provides a unique perspective on the composition of housing wealth inequality, as he tried to understand how housing wealth inequality is distributed by region and age of households. The study investigated a variety of wealth measures like gross housing wealth and net wealth. Henley showed that there was a small growth in housing wealth inequality over the period 1985-91. Notably, Henley points to the growth of inter-region inequality as the main source of growth in inequality.

While Henley's analysis was quite comprehensive, it is not without its flaws. Levin and Price critique the use of BHPS data since the dataset lacks spatial information so true regional analysis is not possible with this dataset.[10] Furthermore, they also critique the lack of available transaction data, resulting in approximate estimates of housing wealth in most cases. They believe that these limitations make regional analysis through the usage of BHPS data impossible.[11] They advise the usage of house transaction data as it can provide price and attribute information while also allowing for extensive geographical coverage as well as more detailed spatial characteristics.[12] They also acknowledge that

[8] Levin E. & Price G., "Measuring Changes in Housing Wealth Inequality", Department for Communities and Local Government, Economics paper 6: Volume 2, 2010
[9] Henley A., "Price Shocks, Windfall Gains and Hours of Work: British Evidence", Oxford Bulletin of Economics and Statistics, 66 (4): 439-456
[10] Levin E. & Price G., "Measuring Changes in Housing Wealth Inequality", Department for Communities and Local Government, Economics paper 6: Volume 2, 2010
[11] Ibid
[12] Ibid

usage of this data, results in a loss of mortgage information which Henley and MacLennan and Tu would argue is a serious oversight, as it discounts the potential mortgage debt that a household might hold and therefore overestimates the overall value of the households wealth.[13]

However, we see Thomas and Dorling use transaction data to examine gross housing wealth differentials in their paper despite these issues.[14] Thomas and Dorling find that the housing markets are responsible for the growth in wealth inequality in Great Britain. The state housing wealth inequality has reached unprecedented levels and is insurmountable for most households through normal means.[15] Although Levin and Price agree with their data choice, they see the focus on net housing wealth instead of gross housing wealth as a flaw in the study.[16] Arguing that at any point in time borrowers are likely to have high gross debt but low net debt since their mortgage is close to maturity.[17] Therefore, rates of outright ownership are a poor measure of debt held and can lead to biased measures of net housing wealth. They also argue that ownership rates suffer as the long run measure of the geographical distribution of inequality, due to the concentration of older, lower debt households in specific areas which can give an impression of spatial inequality which is misleading.[18] The other problem with the Thomas and Dorling paper is that they compare measures between two points in time rather than analysing all the points simultaneously, which can generate biased measures of inequality.

This brings us to the more recent studies published by the Ministry of Housing in November 2010. All the papers examine housing wealth inequality in Britain and how it has changed over time. Chris Hamnett's paper argues that there are significant inequalities in the distribution of regional housing

[13] MacLennan D. & Tu, Y., "Changing Housing Wealth in the UK, 1985 to 1993: Household Patterns and Consequences", Scottish Journal of Political Economy, 45 (4): 447-465
[14] Thomas B. & Dorling D., "Know Your Place Housing: Wealth and Inequality in Great Britain 1980-2003 and beyond", Shelter
[15] Ibid
[16] Levin E., Price G., "Measuring Changes in Housing Wealth Inequality", Department for Communities and Local Government, Economics paper 6: Volume 2, 2010
[17] Ibid
[18] Ibid

wealth in Britain. The paper points to income inequality as being responsible for the rise in wealth inequality, which supports Atkinson's conclusion.[19] The paper continues to point out that housing wealth has historically been a force for wealth equalisation.[20] While it makes sense how income inequality can impact housing wealth inequality, it cannot be the sole determinant for the variation. Thomas and Dorling argued directly against this in their paper, emphasizing that wealth equalisation could not occur through normal income augmentation, like switching to a higher paid job.[21]

This is a common theme amongst papers in the first volume. Pryce's paper suggests that inequalities in housing wealth can have negative effects on access to education, racial segregation and spatial concentrations of poverty.[22] He continues to demonstrate that house prices can reduce the ability of low-income families to take advantage of the best schools. He concludes that housing wealth inequality breeds further housing wealth inequality through the side effects mentioned above.[23] This analysis adds additional layers of complexity to the housing wealth inequality loop we suggested earlier. Furthermore, the paper ends with a suggestion for further analysis of the data on the inter-regional level as well as over a longer period.[24] Our paper is a response to this request.

The Equality Trust's regional inequality briefing notes demonstrate the link in variations between property prices and property wealth in the UK.[25] This suggests that the price of someone's house is the main factor in their amount of property wealth, rather than the number of properties, or the percentage of property owned. This adds further support to the use of variations in property prices as an estimate for housing wealth inequality in England. In addition, the paper shows that London appears to be

[19] Hamnett C., "Housing and Wealth in Britain", Department for Communities and Local Government, Economics paper 6: Volume 1, 2010
[20] Ibid
[21] Thomas B., Dorling D., "Know Your Place Housing: Wealth and Inequality in Great Britain 1980-2003 and beyond", Shelter
[22] Price G., "Has house price inequality risen or fallen since 1996?", Department for Communities and Local Government, Economics paper 6: Volume 1, 2010
[23] Ibid
[24] Ibid
[25] "A Divided Britain? Inequality Within and Between the Regions", The Equality Trust

somewhat of a paradox since it is considered the richest region in the UK but also one that experiences the most severe poverty rates.[26] The main conclusion of the paper, it that inequality in Britain is most apparent between individuals rather than particular regions, which is an interesting counter-argument for the existence of the North-South divide in England.

We see a similar argument employed by Clark and Cummins with regards to North-South housing wealth differentials. They argue that there is a form of natural sorting present, that attracts the less economically successful households to the North with the South experiencing the opposite.[27] Notably, their analysis demonstrated the average housing wealth within regions, which masks the housing wealth inequalities that can be seen within regions. This increases the importance of our measure since it will demonstrate the effects of this type of sorting over a larger range, which will allow us to see how less economically successful households struggle in areas of high economic success.

The study closest to our methodology is the "Delivering Changes in Housing Wealth Inequality" by Levin and Pryce. Their paper attempts to analyse housing wealth in more depth than previous studies. The main goal of the paper is to determine whether housing wealth inequality is a genuine phenomenon or just the side effect of varying datasets and approaches to the topic.[28] They argue that inequality of gross housing wealth has fallen for the period 2000-2006, which seems to be the result of the regular cycle of housing wealth inequality, even so, they do acknowledge that the shift in the distribution has increased the divide between homeowners and renters.[29] We will test whether these findings hold up in the long run by generating similar measures for our period of the dataset (1995-2018). Since we are using the

[26] Ibid
[27] Clark G. & Cummins N., The big sort: The decline of northern England, 1780–2018, https://voxeu.org/article/decline-northern-england-1780-2018
[28] Levin E. & Price G., "Measuring Changes in Housing Wealth Inequality", Department for Communities and Local Government, Economics paper 6: Volume 2, 2010, 12
[29] Levin E. & Price G., "Measuring Changes in Housing Wealth Inequality", Department for Communities and Local Government, Economics paper 6: Volume 2, 2010, 42

same primary source, this will be a good way to check the validity of our estimates.

Research Design/Methodology

This paper seeks out to construct a time series for regional housing wealth inequality in England over the period of 1973-2018. We intend to use this time series to demonstrate how the housing wealth gap between regions has evolved over the period. Our main measure for inequality will be the standard deviation between regional means of housing prices. We also include more conventional inequality measures, like the Gini and Atkinson coefficients, after 1995 when the data available is more comprehensive. These values will be calculated using the ineq function from the r package "ineq", we can generate both measures by setting the type parameter to our desired measure of inequality.[30] Furthermore, for the period 1995-2018, we also construct similar measures for inter-regional housing wealth inequality.

For years before 1995, we intend to use and contrast two major sources of housing transactions. Our first and the most relevant dataset, the Nationwide Housing Price Index (NHPI). This dataset begins in 1973 and provides quarterly seasonally adjusted estimates of mean housing prices of each region in the UK. The NHPI is derived from Nationwide lending data for properties at the post-survey approval stage.[31] The index is mix-adjusted which reduces the influence of changes in property mix, generating less biased mean estimates.[32] Our second dataset is the Halifax Housing Price Index (HHPI). The HHPI starts in 1983,

[30] Detailed breakdown of package and function at https://cran.r-project.org/web/packages/ineq/ineq.pdf
[31] Nationwide Housing Price Index, 2018, https://www.nationwide.co.uk/about/house-price-index/methodology#tab:Methodology
[32] Ibid

10 years later than the NHPI, with a similar methodology to the NHPI. Notably, the HHPI values tend to be higher on average over the period than NHPI. Both series demonstrate similar trends and are highly correlated with one and another as well as the Land Registry data. Although we noticed that the NHPI measure seems to display a much higher degree of correlation with the Land Registry estimates than the HHPI. Due to this as well as an additional 10 years of data, we will be using the NHPI as our main measure of housing wealth inequality before 1995.

The measure will only account for households that own property. While measures could have been made to take account of the renters in the calculations after 1995, little could be done before 1995 as the data is aggregated and cannot be easily adjusted. Although this analysis does have implications for the renter as well as property-owning households since an increase in housing wealth inequality can imply that the proportion of affordable housing is falling, therefore it can imply rising rents for renters which in turn reduces their disposable incomes and as a result their welfare. A significant rise in housing wealth inequality may even result in crowding out effect of certain renters, reducing their social mobility.

We also underestimate the top end of the distribution since we assume that most households own a single property, which is not the case for the most affluent households. Although the most affluent households should be considered an anomaly in our analysis, since they usually hold their wealth in a number physical and financial assets, therefore movements in housing wealth inequality are likely to have minimal impact on the overall welfare of these households. Therefore, this analysis is not representative of the tails of the distribution, but rather the middle.

We decided to analyse the standard deviation, rather than the mean since it lets us compare the differences in housing wealth held in a given period. The focus on the standard deviation of regional means allows us to get a clearer representation of shifts in the distribution since we are interested in the

overall change in the middle of the distribution and not the tails. Although an argument could be made for the usage of medians instead of means, our data before 1995 is too limited to accommodate this type of measure. Therefore, to keep our measures consistent throughout the whole period, we will use the standard deviation of regional means as our main measure. This measure will be supplemented with several other measures after 1995 when the price paid data from the Land Registry becomes available. Also, the Gini and Atkinson coefficients will not be as effective at demonstrating the change in wealth during the period since we discounted the tails of our distribution. Future studies could investigate including these tails in the calculation, while also analysing them separately to see whether our assumption about the characteristics of the affluent households holds true.

Another major flaw with our measure is that it assumes households must own their property outright for it to be considered as actual wealth, which is not the case for most of the population that finance their property purchases through mortgages and hence incur a large amount of debt during the purchase. This returns to the argument that was proposed by Pryce and whether it is right to estimate wealth through gross or net measures.[33] However, due to the length of our study, we can assume that mortgages made at the start of our time series (1973) have been fully paid off by 2018, and as a result, the increase in property values during the period has been added to the household's wealth stock. While we are looking at the current transactions in each year of our time series, we only consider them to represent the current state of the housing market rather than actual housing wealth inequality. Since housing is a relatively illiquid asset, we cannot infer that transactions are a perfect representation of actual housing wealth held. Therefore, our measures in any given time instance should only be interpreted as estimates of possible wealth held. Even so, this does not invalidate our main goal of analysing the trend of housing wealth inequality over the period. Indeed, it is safe to assume that most

[33] Levin E. & Price G., "Measuring Changes in Housing Wealth Inequality", Department for Communities and Local Government, Economics paper 6: Volume 2, 2010

properties would be sold at least once during our period of study. Hence, we can assume that our measure is more robust in the long run which makes it valid for our analysis. The data supports our critique, as the composition of housing transactions in any given year underrepresented the actual proportion of detached properties in England.[34] This is expected as detached properties have the highest mean value in our dataset and are therefore more illiquid. This quality does reduce the accuracy of our yearly estimates for the upper tail of the distribution, therefore we can assume that our yearly calculations are underestimating housing wealth inequality in any given year. However, when we consider the long run composition, the transaction data comes remarkably close to official estimates validating our initial assumption and making our measure valid. A consideration for follow up research is to mix adjust the values to reduce the influence of bias generated by composition of yearly transaction data.

Primary Source

The main dataset for our analysis is the Land Registry dataset, which is freely available for download.[35] Specifically, we will be using the price paid dataset. This is an extremely comprehensive dataset with nearly 24 million property transactions in England and Wales, that contains all the standard and additional price paid transactions received by HM Land Registry since 1st of January 1995.[36] This dataset is updated monthly making it an extremely valuable resource for analysing current as well as historic changes in housing prices. The dataset is quite a large file, 3.7GB on average. The share size of this

[34] See Table 1
[35] See bibliography for link
[36] HM Land Registry, 2018, https://www.gov.uk/government/statistical-data-sets/price-paid-data-downloads

dataset has provided several challenges in our analysis, as the dataset is so vast it cannot be simply opened and processed in Excel, therefore we had to switch to R which is better equipped for handling files of this magnitude. However, R does not magically fix the scalability problem of this dataset. We experienced significant waiting times for operations like na.omit(), which was used to remove all observations with NA values, in the data cleaning stage of our analysis. Notably, this analysis should not be run on machines that have less than 16GB of RAM, as certain data manipulation cause RAM utilisation to exceed 12GB on certain occasions. Even with a powerful machine, care needs to be taken with code efficiency. The inefficient code can slow down the analysis considerably at best and result in R crashing in the worst case.

We have found that removal of unnecessary columns sped up our analysis greatly, it allows us to use less RAM for storing the data, while also reducing processing time due to a significant reduction of data processed with each operation. We managed to reduce the overall size of our dataset from 4GB to 1.6GB. This is heavily encouraged for anyone interested in exploring this dataset as it will allow for more complex analysis as well as increase margin for error with the code. It should be noted that this dataset is available in monthly and yearly instances, these files are considerably smaller and are recommended for anyone planning to analyse this dataset without the need for all the data.

The dataset is quite comprehensive, in addition to the nominal value of each transaction we also see a few other useful characteristics. There are 16 columns in this dataset, one for each characteristic. Notably, most of this information is intended to describe the exact location of each transaction, for example, the PAON category is referring to the specific house number/name.[37] This kind of granularity in the dataset allows for precise spatial analysis, while this data is quite valuable we are not concerned with this level of precision in our analysis. Therefore, we can remove most of the categories related to

[37] Ibid

location, leaving only the county category. We left county since this is the most macro measure of the location provided, notice how the land registry is missing a regional classification. This is problematic since our analysis is focused on regional inequality. As a result, we had to add an additional category for the region of each transaction in our dataset. We focus on regional inequality since it matches the precision of the NHPI and the HHPI allowing for easier comparisons. We intended to match the 10 regions we were given in the NHPI/HHPI, however "Outer Metropolitan" was quite a vague classification which we omitted in our classification.[38] We also decided to remove Wales from our dataset since this is outside the scope of this study. While significant care was taken to avoid unnecessary loss of data, we still lost over 2 million observations during our region classification. The removal of Wales is responsible for most of this loss. However, a small portion of the data loss is due to label inconsistencies like the classification of a city as a county. The period is also responsible for label inconsistencies since there were several counties that changed names or were absorbed into other counties. While this was accounted for during the classification, some data loss was still inevitable. Although this loss is unlikely to have significant effects on our analysis since we still have nearly 21 million observations available. Even so, the precision of our estimates still suffers as a result. Our final dataset had the following variables: Price, Type of Property, Date of Transfer, County and Region. This improved the performance of our analysis dramatically allowing us to comprehensively explore the dataset.

After further exploration, we removed all transactions that were classed as classed us "other" from our analysis, due to a few reasons. After 2011, these transactions reached extreme values, in excess of £500 million, quite consistently and in relatively large quantities, as a result, we saw a significant increase in the standard deviations after 2011, with peaks of £2 million. We assume that these transactions are not representative of regular household transactions. In addition, the other category is not well defined in

[38] NHPI and HHPI datasets

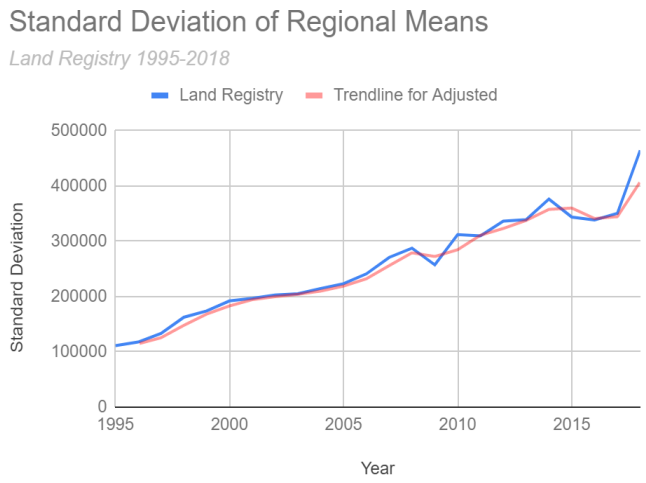
the documentation for the dataset, making an accurate analysis of this category quite difficult. We can safely assume that the other category might refer to actual transactions of land rather than built property while this is valuable information, the complications it generates in our analysis are too great to justify keeping the category in our analysis.

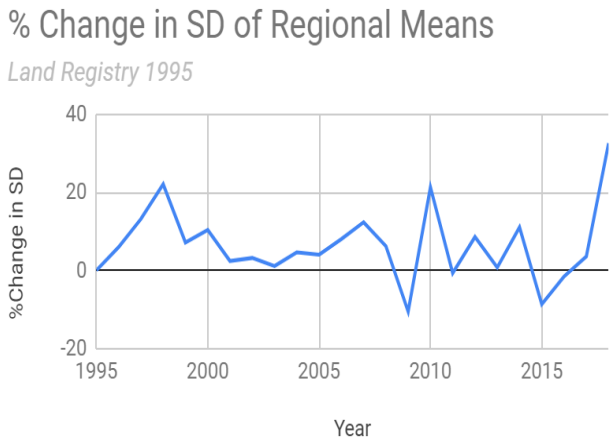
Another factor that we should consider in our analysis is inflation over the period. Since all of our data is nominal, we have to adjust it for inflation. This applies to the NHPI/HHPI datasets as well as the land registry. We used the RPI values for the UK from the ONS (ONS RPI Inflation, 2018). RPI was chosen over CPI due to CPI being a relatively young measure and therefore not having data that goes back far enough. Another advantage of using RPI is that it accounts for inflation in housing which CPI does not.

Analysis

Moving on to our analysis of regional housing wealth inequality in England. Figure 1 demonstrates the change in the standard deviation of regional means. The trend line indicates a clear real increase over the period implying that the regional inequality has increased since 1995. We can also see evidence of the 2008 financial crisis in Figure 2, where we can see a dramatic fall in the growth rate. This was probably because the financial crisis made the purchase higher-end properties less attractive during the period. Although we also see a similar dip in 2015 which was the year of the Brexit referendum. The fall in 2008 is interesting as we saw mean values within every region besides London fall, which should indicate a rise in regional housing inequality, not a fall. So, despite the financial crisis, the average value of a property in London increased. Although the number of transactions fell dramatically in 2008, falling 49% from 2007. While 2015 saw a slight increase in property transactions with 1.6% more transactions than in 2014. In addition to this, the regional mean values increased in every region during this year

implying that a catch up among regions during this year. Indeed, the data shows that the South of England (excluding London) and Midlands had a significantly higher growth rate than London and the North.



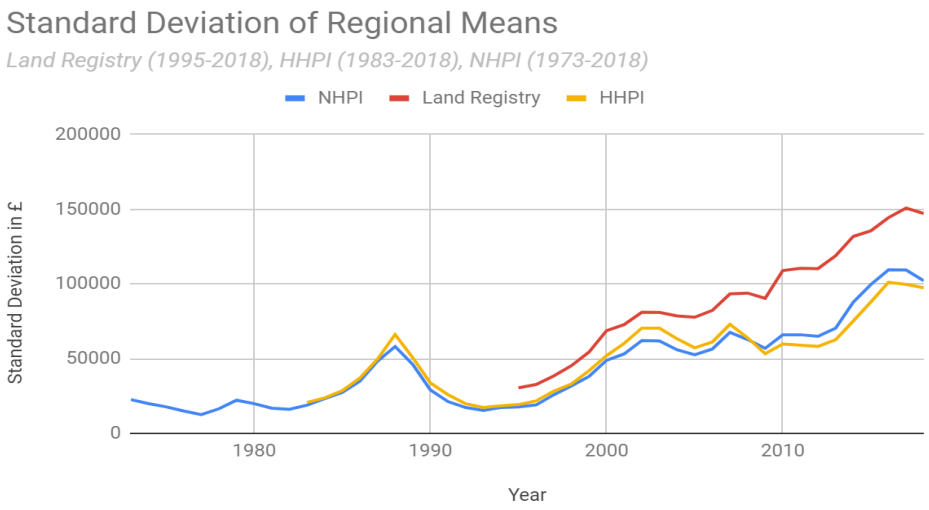


Next, we examine how our dataset performs when compared to our secondary indexes. Figure 3 shows our land registry values alongside our estimates calculated from the NHPI and HHPI datasets. We can see that NHPI and HHPI are highly correlated (0.972), while also being highly correlated with the land registry estimates (0.972 and 0.914 respectively). We can see that the land registry estimates are significantly higher than their NHPI/HHPI counterparts, however, this is expected since NHPI/HHPI are based on mortgage data which implies they are discounting larger transactions that are made without mortgages.

While the NHPI/HHPI may provide underestimates of regional housing wealth inequality, the high correlation with the land registry data implies that they are good indicators of trend. As a result, we can use these two indexes to examine trends in periods before the price-paid data. This allows us to extend the start of our time series from 1995 to 1973. With this extension, we can see that the rise in regional

housing wealth inequality is quite a recent phenomenon. Besides a short period in the late eighties, we can see that regional housing wealth inequality has been quite stable historically until 1995. After 1995, we can see a consistent increase in the standard deviation of regional means.

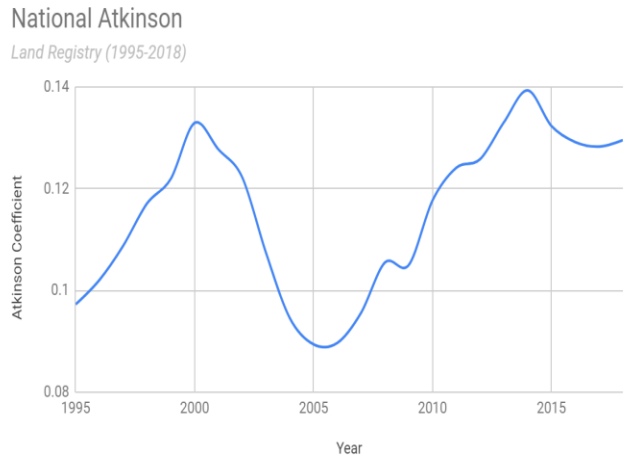
The increase in inequality that we see in the late eighties is most likely the product of the 1980s housing boom which experienced a speculative bubble towards the end of its lifecycle (Ball M., 1994). While the bubble was quite extreme with annual growth in housing reaching record levels, it was short-lived resulting in the crash of housing prices as well as overall housing wealth in the country. What we see in

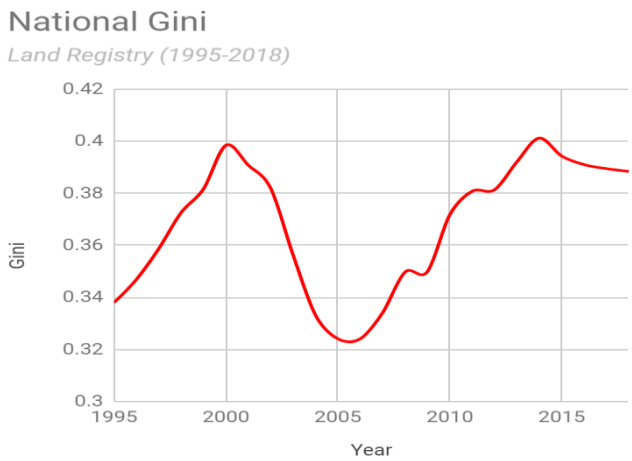


the early nineties is the deflation of the speculative bubble. Although this raises further questions, why did we not see a similar deflation in 2008? While we did see a fall in 2008, it was insignificant compared to the crash we see in the late eighties. From this we can interpret that our measure of regional housing wealth inequality seems to have a cyclical element, however, it does not seem that be evident after 2000 where we consistent increases in our measure. From the data presented so far, we can infer that recessions tend to have an equalising effect on regional housing wealth inequality. Although this is more applicable before the twenty-first century.

Next, we will compare our findings to another more conventional measure of inequality, the Gini

coefficient. We also calculated an Atkinson coefficient with ε = 0.5, while we could have included additional values for ε it does not seem necessary for this comparison. From the figure below, we can see say that the Gini and Atkinson coefficients are highly correlated. Furthermore, both measures exhibit cyclical properties during the period. These findings are consistent with Levin and Pryce's original findings in 2010.[39] Both the Gini and Atkinson coefficients have peaks in 2000 and 2013, while also being at their lowest during 2005. The fall in these coefficients might represent the ease of obtaining a large mortgage during the early 2000s which could have allowed more individuals to purchase homes of higher value than in other periods. We can see that the contraction of easy credit, due to the financial crisis of 2008, seems to have made the situation worse which would explain the increase in inequality we saw after 2008. Both coefficients suddenly fell in 2008, this behaviour is similar to our estimates of regional standard deviations in housing prices, implying that the 2008 crisis had a minor equalizing effect on regional housing wealth inequality. Furthermore, we can see that all of our measures rose in the period 1995-2000, therefore we can assume that the housing crash of the late eighties also helped to deflate the rise of regional housing wealth inequality during the housing boom. Notably, it has been slowly falling since 2013 although nothing like the changes we saw in the early 2000s.

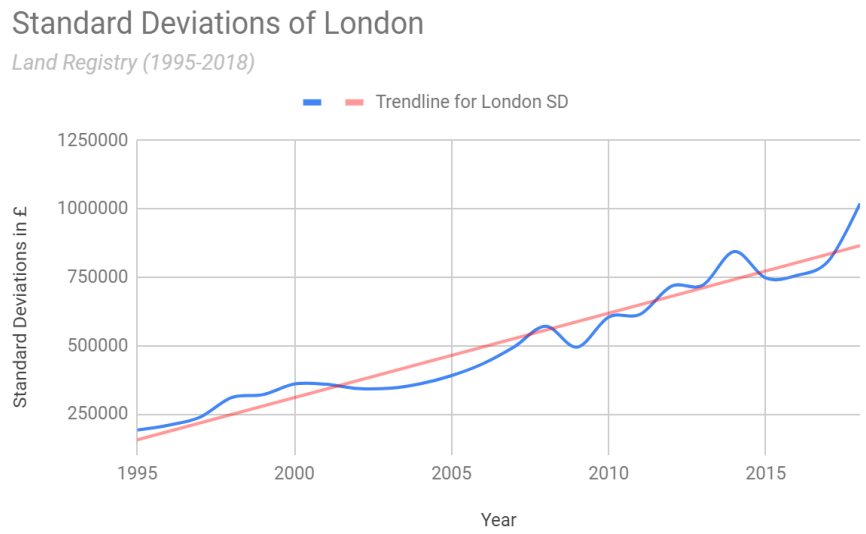




[39] Levin E. & Price G., "Measuring Changes in Housing Wealth Inequality", Department for Communities and Local Government, Economics paper 6: Volume 2, 2010

While these measures give us an idea of how housing wealth inequality evolved between regions, they do not take advantage of the granular precision of our dataset. So, it makes sense to take advantage of our dataset and extend the analysis within regions.

London



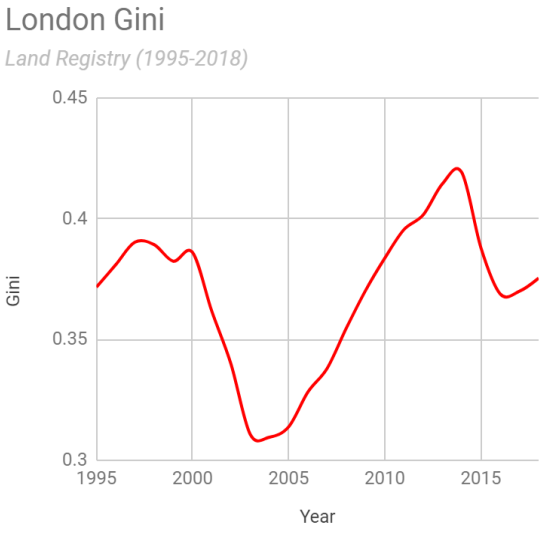
London will be the first region we analyse. London is an interesting region since it consistently has the largest number of transactions while also boasting the largest mean value for housing. This is to be expected of one of the world largest financial centers however it does not undermine the housing crisis that results from these accomplishments.[40] While the housing crisis poses a serious problem, it perfectly demonstrates what are the side effects of extreme housing wealth inequality. Furthermore, the housing crisis is the result of London's housing bubble however this bubble has yet to burst as we shall demonstrate. The figure below demonstrates the increasing wealth disparities in London. As we can see from the plot, the housing wealth inequality in London has increased dramatically over the period. Our

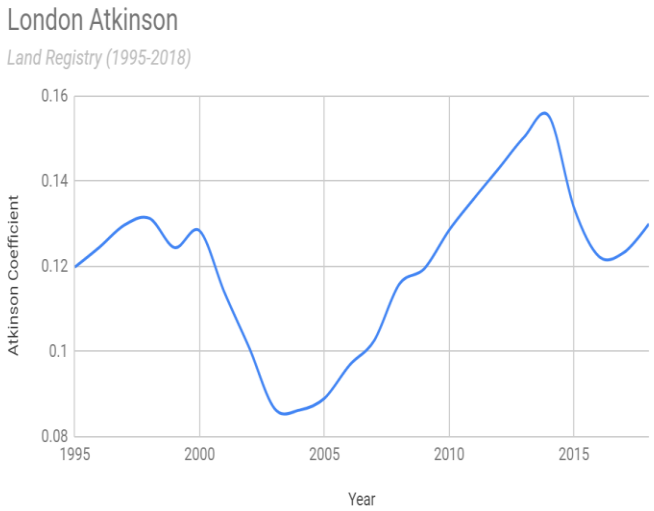
[40] Travers T., Sims S., Bosetti N., "Housing Inequality in London", Center for London, 2016: 16

initial values were quite close to the other regions at the time, however as time progressed the London housing bubble has elevated housing wealth inequality in London to unsustainable levels. While we can see slight deviations from the trend line in 2008 and 2015, those are the outliers of this time series. As a result, we suspect the majority of regional housing wealth inequality can be explained by this trend. Travers, Sims, and Bosetti showed that the majority of wealth differentials between Great Britain and London is due to housing wealth.[41]

Now we should examine how London performs in terms of the Gini and Atkinson coefficients. The figure below shows that the London Gini and Atkinson coefficients are cyclical and also correlated with the National estimates. London actually has the highest correlation with the national estimates with an average value of 0.7. This is expected since London is responsible for a large portion of all housing transactions. Although this implies that our national measure is slightly unbalanced due to London and hence can skew the results. This is a notable flaw which should be addressed in later studies.

Despite this, we can some unique trends present in our London Gini/Atkinson estimates. While we still see a fall in the early 2000s, as we did in the national estimates, we see that the lowest point is before





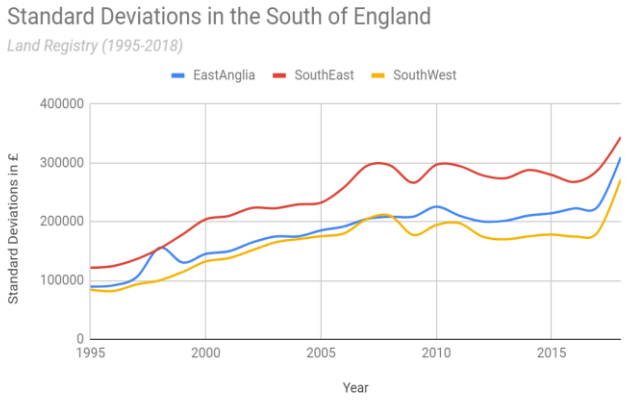
[41] Travers T., Sims S., Bosetti N., "Housing Inequality in London", Center for London, 2016: 33

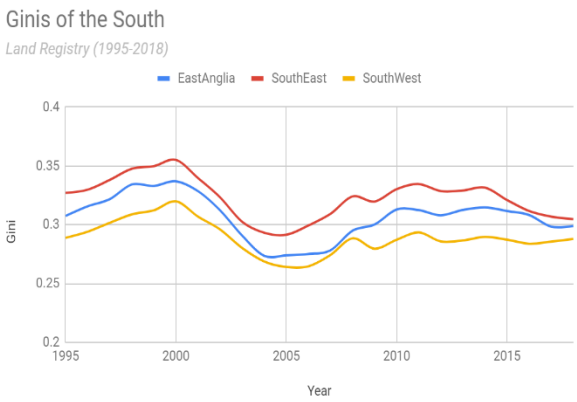
2005, this fall can also be seen in the standard deviations however the effect is less pronounced due to the scale of later transactions. Furthermore, we can see that the 2008 financial crisis is not visible in our data. While the standard deviations and the fall in the total number of transactions capture this effect, the Gini/Atkinson estimates are not affected. This could imply that property was seen as a safe asset during a period of uncertainty, therefore adding more air to the London housing bubble. This explanation makes sense, since we saw both the Gini/Atkinson estimates and the mean in London rise during 2008, it could signal a crowding out effect in London housing, where investors speculate on London housing, which drives the prices up, in turn making the properties less attainable for normal households, crowding them out of the market.

However, the most interesting trend is the sudden fall in the Gini/Atkinson coefficients after 2013, when we see a dramatic equalisation in terms of those measures. In addition, we see the standard deviations fall alongside them. While this may seem like a turning point for the better, it is not. When we look at the mean value of a property transaction in London, we see that prices increased by more than 8% per year from 2012 to 2014. This could be seen as the peak of the London housing bubble since the mean value has continued to climb but at a more reasonable rate.

The South (excluding London)

Next, we will look at regions closest to London, the South of England. This section will cover multiple regions rather than focusing on just one at a time. This is mostly due to most regions in the South of England (excluding London) being quite similar and therefore exhibiting similar trends to their neighbours. This is shown in the figures below.

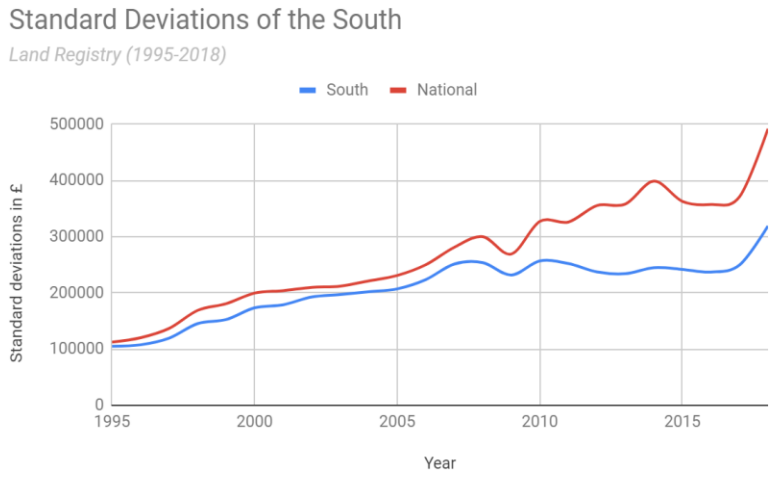




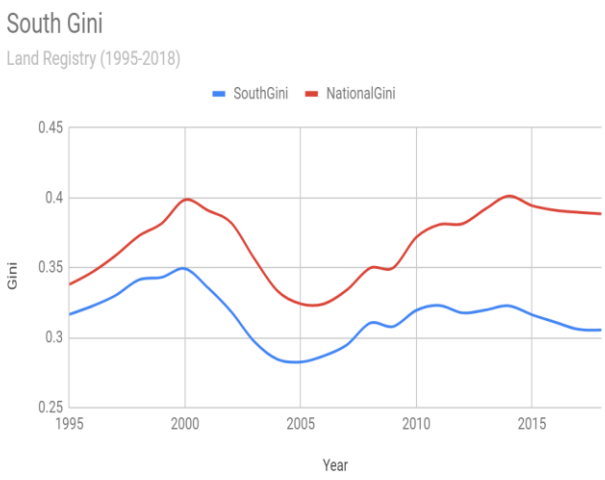
While we can see that there are some minor differences, the main being seen in East Anglia in 1998, the overall trend is very similar. The three southern regions are highly correlated with each other with correlation coefficients of 0.9. We can see that the South West of England has the lowest housing wealth inequality in the South using both of our measures. While the South East of England seems to be the most unequal during the period. It seems like proximity to London has a noticeable effect on housing wealth inequality within the region. Since the trends are extremely similar we will analyse them in aggregate by classifying all the transactions in South East, South West, and East Anglia as transactions in the South of England. The plots for the South of England are shown below with contrast to the national values.

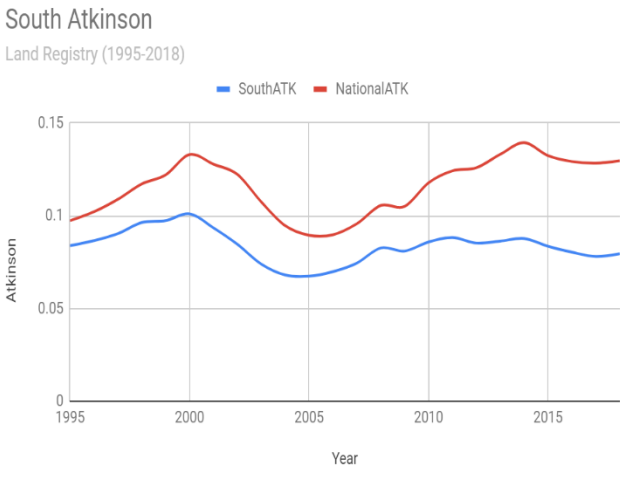
From the plot, we can see those standard deviations of the South are highly correlated with deviations of regional means until 2010. At that point we see the national values increase dramatically, almost doubling the South's values by the end of the period. Most likely, this is the consequence of London's

housing bubble which is pushing up the national average due to speculation. Although, this might not be the case since we see the South follow a similar trend as the national measure after the slowdown of London's property market in 2015. This could imply that overall speculation shifted from London property to properties in the South of England that are near London. Although that is slightly misinformed since all regions in the South experienced this increase in deviations us we have shown.



Next, we should look at how the Gini/Atkinson coefficients of South changed during this period. They are shown in the plots below.



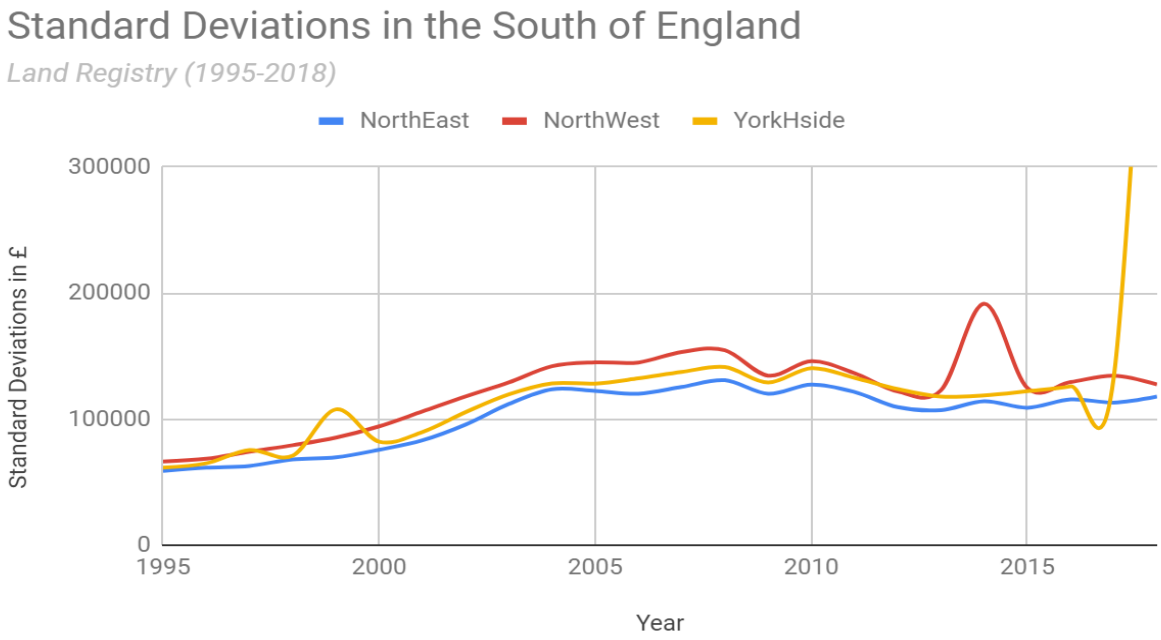


From these plots, we can see that the South's Gini/Atkinson values are also cyclical and highly correlated

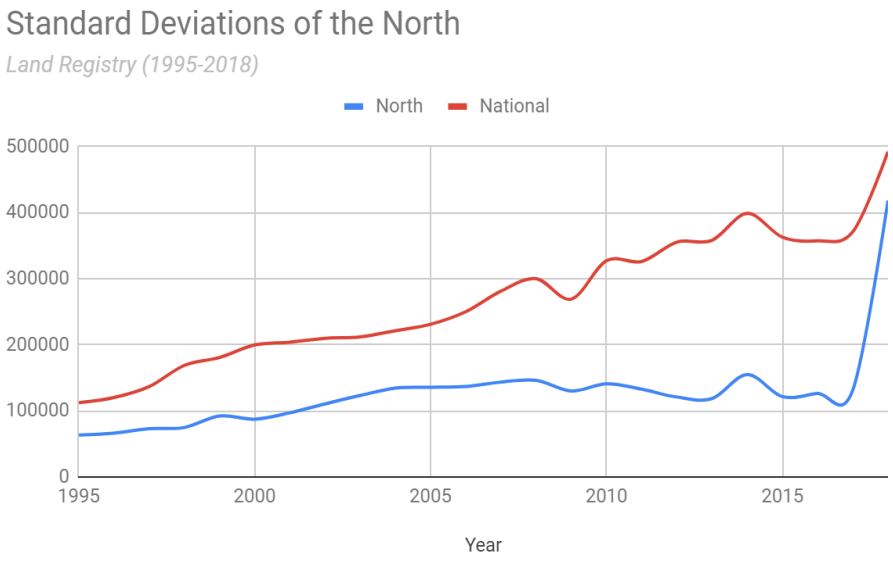
with our national estimates. There is really nothing additional we can say about housing inequality in the South since it follows the same path as the national estimates. We can infer that the South of England follows the national trend and has lower housing wealth inequality than England as a whole.

The North

Next, we will analyse the performance of the North of England during this period. We intend to do another aggregate analysis as we did for the South of England. Again, this is due to some regions having extremely high correlations of 0.9 between themselves. We will classify the following regions as the North: North East, North West and Yorkshire and the Humber. Notice how we have excluded the West and East Midlands from our classification, this is due to their relatively low correlation values with regions in the North. We also considered including them in our South category however we encountered the same problem. Furthermore, we have found that the West and East Midlands are highly correlated with one another, which justifies analysing them in a separate group.

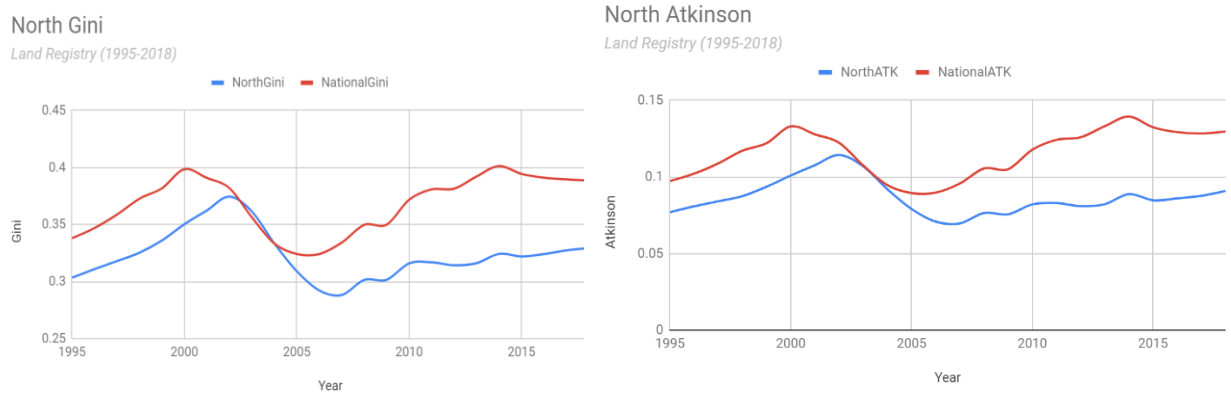


So, from the plot above, we can see that the standard deviations of Northern regions are quite similar, however, there are some interesting outliers. Yorkshire and the Humber seem to experience a shock in the late nineties which results in a temporary rise in housing wealth inequality, although this was quite minor. In 2013, we see a temporary shock in the North West of England that nearly doubled the standard deviation of the region. This could have been caused by possible investment in the North West as our dataset does not distinguish between civilian and business purchases. The most striking value seems to be our 2018 measure of Yorkshire and the Humber which seems extremely out of place given the time series we constructed. Most likely, this is an outlier similar to the 2013 shock in the North West, which implies that it is not representative of housing wealth inequality and should be omitted from the analysis. This is why we should use our aggregated analysis since it will correct the majority of these outliers and give us a more accurate trend.



Although our aggregated measure has reduced the outliers, it has not eliminated them. We can see from our graph that the outlier from 2018 in Yorkshire and Humber has a significant influence on our measure. Although a rapid increase like this cannot be representative of changes in the distribution of

housing wealth in the North, so we should conduct our analysis without it. Over the period, we can see that there is relatively little change in the standard deviation of property transactions in the North. Implying that housing wealth inequality is quite stable in the region if we discount the outlier.



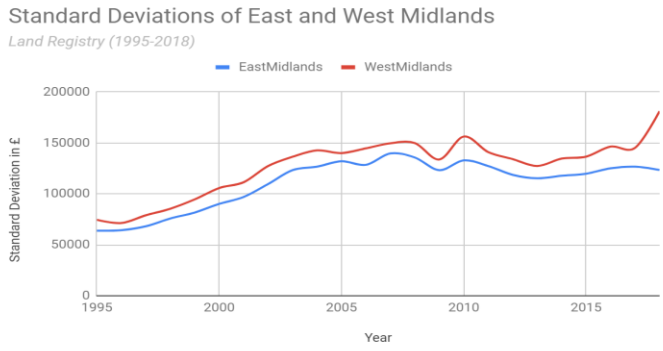
From these plots, we can see that the North's Gini/Atkinson values are also cyclical however the trend compares differently to national estimates. We can see that initially the North's Gini/Atkinson estimates lagged behind the national trend. However, after 2005 this was no longer the case. Notably, the Gini/Atkinson values tend to be higher in the North than the South during the period. This implies that the distribution of housing wealth is better balanced in the South than in the North, this is despite the South having a much higher standard deviation of property prices during the period. Although it is possible that the non-household transactions might be inflating our measure of housing wealth inequality in the North.

The Midlands

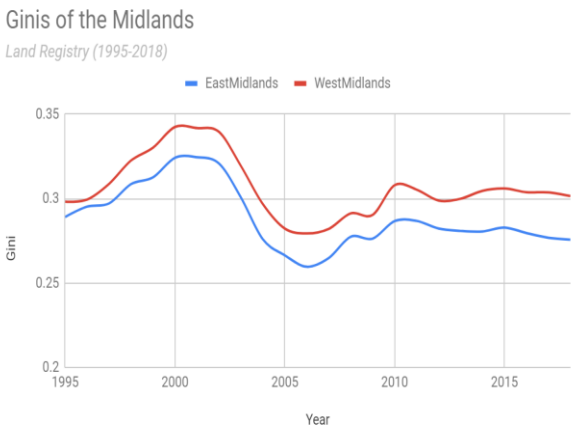
As we mentioned before, the transactions from the West and East Midlands were grouped separately from the other regions due to the high correlation they shared amongst themselves. This also makes our analysis more unique as very few of regional inequality classify the Midlands as a separate macro-region. While we will add our thoughts on the North-South divide in England, which is a common topic of

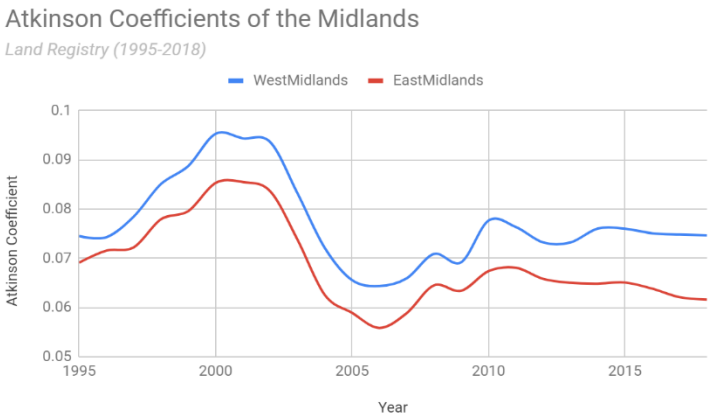
debate in studies of regional inequality in England, we will also add our own take on the debate with the help of this dataset.

Initially, we will show how similar these two regions are, which should justify our aggregate analysis. The first plot demonstrates how similar the trends from these two regions are. We can see that the West has more variation during the period, implying higher housing wealth inequality than the East. We can also see that variations as a whole have increased over the period in both regions which is quite similar to the other regions we analysed. We should also note that both regions experienced a decline in variation during the 2008 financial crisis, although the West experienced a much sharper decline in variations, followed by a generous increase in 2010. While the East did experience a similar trend, the impact was less pronounced. Furthermore, we can see that the West starts to deviate around 2018, which quite similar to the deviation we saw in Yorkshire and Humber, although not as drastic.

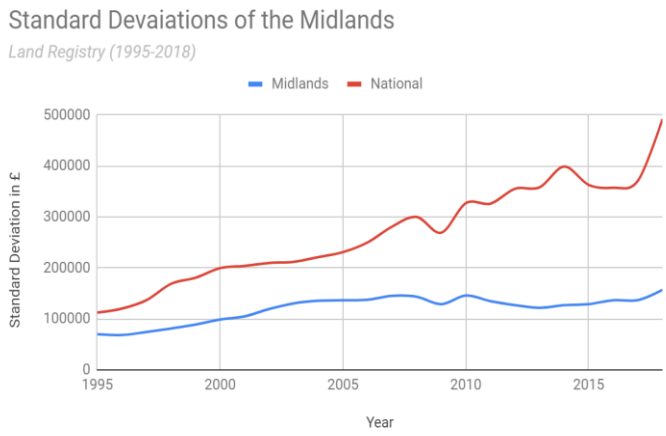


Next, we turn our attention to the Gini/Atkinson estimates for the two regions. Again, we can see that both Gini and Atkinson estimates are highly correlated for these two regions, giving aggregation of these values more validity. Furthermore, the measures confirm that the West Midlands has slightly higher housing wealth inequality over the period than the East.

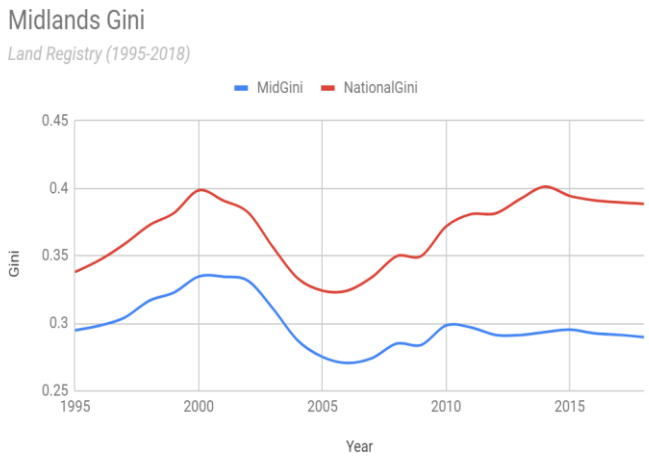


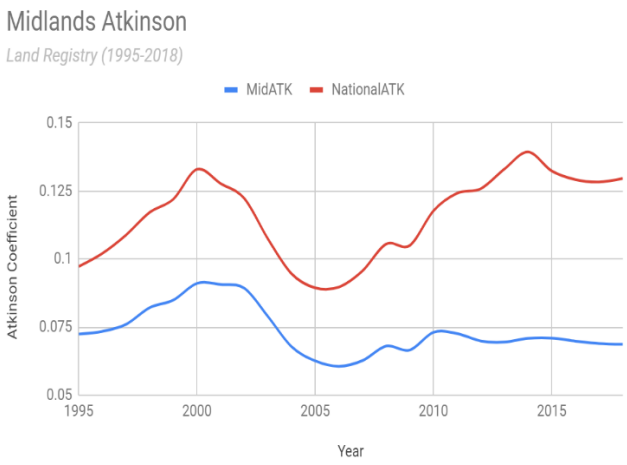


Now we can look at how the Midlands as a whole performed during the period compared to our national estimates. While the Midlands did experience a modest increase in standard deviation during the period, this is undermined by the dramatic increase we saw in England as a whole. According to our estimates, the Midlands has been quite consistent. We can still see the effects of the 2008 financial crisis in our estimates, although these are quite minor compared to other regions. Notably, this trend is quite similar to the trend we saw in the North before 2018. Since both regions experienced little change in terms of standard deviations, it is safe to assume that the majority of the increase is coming from the South and London, although this might be the consequence of generally higher property prices in those regions. So, we should compare the Gini/Atkinson estimates to get a clearer picture.



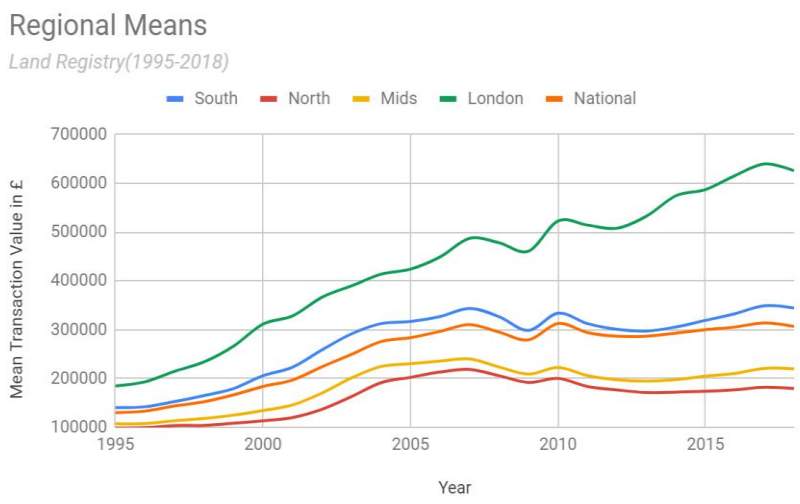
We see similar trends in the Gini/Atkinson values as we saw before. While the values still appear to be cyclical, the changes over the period are quite minor compared to other regions. After 2010, we can actually see that all of our measures of housing wealth inequality have experienced little change. This could imply that housing wealth inequality has been almost constant in the Midlands for the past 8 years, which is quite shocking revelations when we contrast it to our findings.



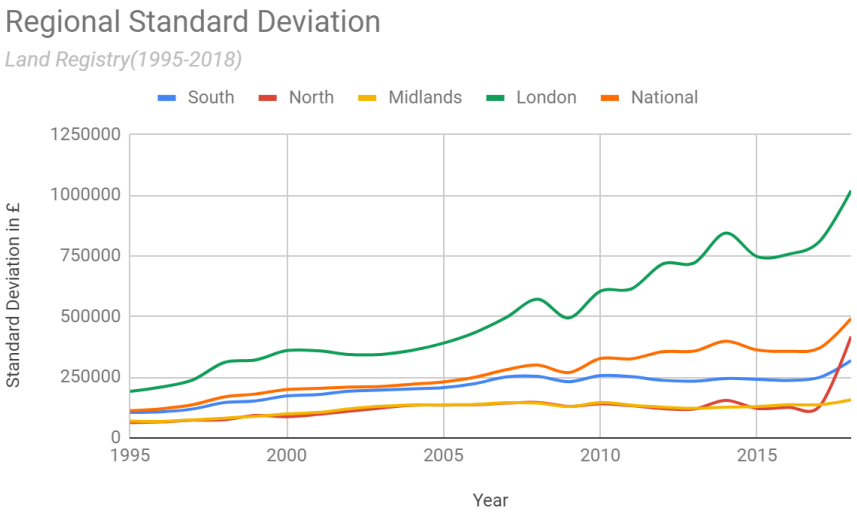


All Regions

Finally, we will analyse all of the regions together, since this will make the causality more apparent. Initially, we will plot the yearly mean house prices for each region during the period, since we omitted this from our region analysis. The mean is a helpful measure since it allows us to see how house prices have evolved over the period, although mean values tell us little about the overall distribution of housing wealth we can still use them to get a better understanding of England's property market.

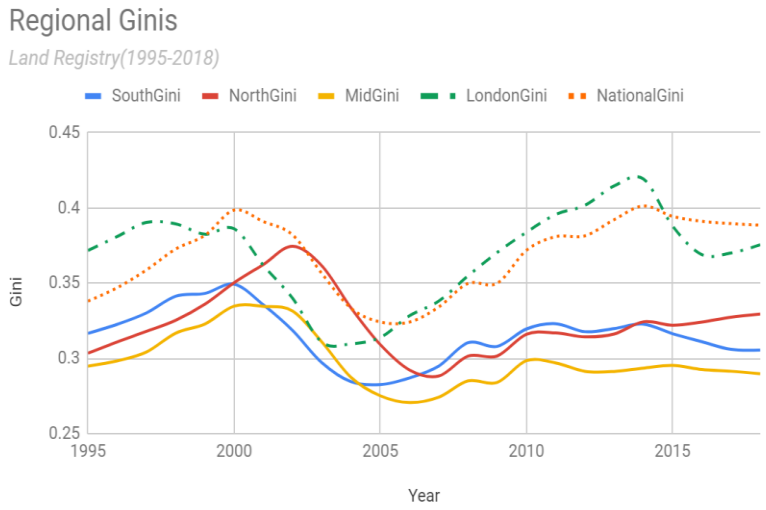


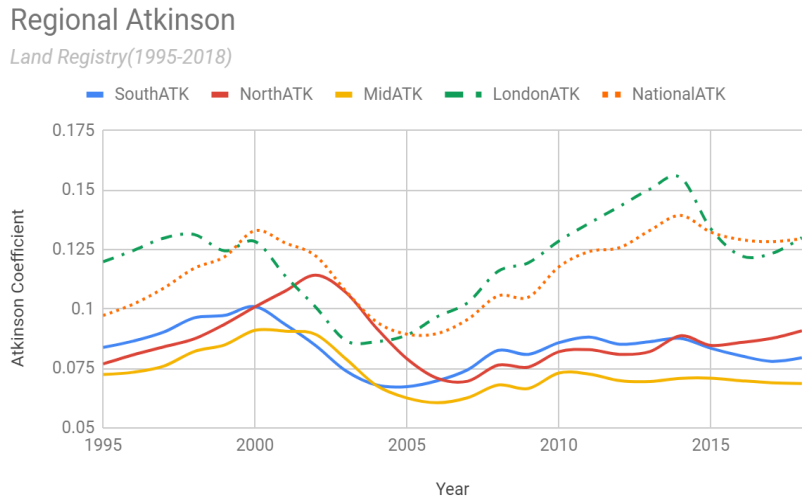
We can see that there is quite a disparity in mean transaction value in London compared to the rest of the country. Furthermore, we can see that the means in the South are also above the national average, although not to the extremes seen in London. We can also see that the means for the South, North, and Midlands are highly correlated, while London seems to be the anomaly in this time series further demonstrating the influence of London's housing bubble. Although the impact of the 2008 financial crisis is still apparent in our time series.



Our standard deviation plot reveals a similar story, we can see that London is responsible for the

majority of the increase in national standard deviations after 2010. Our main outlier in this time series still remains the North in 2018. The only region that has a similar response in London, however, that itself is an anomaly. The plot also demonstrates how closely correlated the North and Midlands really are until 2018. While this outlier is an interesting anomaly, we believe that it does not assist us in answering our question of how regional housing wealth inequality has developed over the period.





The regional Gini/Atkinson values show that these measures are cyclical in nature. These findings

support Levin and Pryce's hypothesis.[42] However, this reduces their viability to demonstrate housing wealth inequality trends in the long run since it is harder to deduce changes in the long run. Despite this, all our measures point towards London as the main culprit for such large deviations in our national estimates. During the period, we can see London has the highest Gini/Atkinson measure, as well as standard deviation and mean of all regions. When we remove London from our comparison, we can see that housing wealth inequality is quite balanced in most regions. Although despite this, we can still see the presence of a divide between the North and South, since the South has significantly higher mean and standard deviation of transactions during the period.

Conclusion

We have established that housing wealth inequality in terms of Gini/Atkinson values tends to be cyclical, which validates the hypothesis of Levin and Price. This property holds even at the region level, with regions being highly correlated to their neighbours. This lets us determine 4 unique macro-regions for our analysis: London, North, South, and Midlands. Our analysis showed that the Midlands was the most stable region over the period, with the smallest shifts in distribution, as well as Gini/Atkinson, estimates after 2010. While London proved to be the most volatile, with standard deviations usually double the national average, and also the highest Gini/Atkinson estimates on average. We actually see that the

[42] Levin E. & Price G., "Measuring Changes in Housing Wealth Inequality", Department for Communities and Local Government, Economics paper 6: Volume 2, 2010

Gini/Atkinson estimates were quite stable in all regions besides London after 2006. Furthermore, the trend in housing inequality during the period 2000-2006 that was described by Levin and Price, can be seen in our time series. Even so, we should mention that the measures are more stable in periods that follow.

Our estimates also confirm the claims made by the Equality Trust, since we can see that London is clearly the richest regions, with the highest mean housing values over the period, but also the most unequal, with the highest Gini/Atkinson estimates on average. We should also mention that London's distribution, mean and standard deviation, shifted dramatically during the period. This implies that the housing wealth gap between renters and homeowners has widened significantly, which in turn reduces the likelihood of renters transitioning to homeowners in London. Although, this trend could also be an example of the sorting effect proposed by Clark and Cummins since the less economically successful households (renters) are being pushed out of London and towards other cheaper regions while the economically successful households (homeowners) reap the benefits of rising house prices.

While the economic intuition of that logic seems quite intuitive, that is far from the truth. The reality is that the majority of economic activity in England is concentrated in London.[43] Therefore, implying that the majority of people work in London and have no alternative to working in London. The renters are forced into a never-ending poverty cycle of rising rents and falling disposable income. Since the renters spend more money on rent, due to rising rents, they save less which reduces their ability to build up enough money to pay for a mortgage deposit and escape the poverty cycle. The London housing crisis is a serious problem which is responsible for a large portion of regional housing wealth inequality.

To conclude we have seen a significant shift in the distribution of regional housing wealth since 1995,

[43] Roses J. & Wolf N., The Return of Regional Inequality: European Regions between 1900 and 2010, Economic History Working Papers

unlike periods before this trend does not appear to be cyclical. We have seen the overall real value of the distribution increase since 1995. While the Gini/Atkinson estimates show that the changes in the composition of property transactions are cyclical, we cannot dismiss the underlying trend present here. Regional housing wealth inequality has increased and the wealth gap between renter and homeowner has become almost insurmountable in the South of England. This trend requires drastic action, as we can see that the symptoms of the London housing crisis are also spreading to its neighboring regions, as people are pushed out of London and forced to commute from nearby regions in the South.

As a final note, our analysis is not intended to describe the causes of this trend. Instead, it is meant to construct the trend and speculate on potential causes. The purpose of this speculation is to encourage more research within the area. Potential directions for future research could potentially include testing the causality of our claims. Looking into whether London's housing crisis is spreading beyond London and how it is affecting the infected regions sounds like a promising topic. Furthermore, our estimates could be refined by mix adjusting our standard deviation values to get a more accurate estimate of housing wealth inequality held. While we assume that our estimates underestimate the yearly deviations in housing inequality, it would be interesting to see if this assumption is correct. Finally, the price paid HM Land Registry dataset has several other potential variables that can be used to assess the impact of certain characteristic on regional housing wealth inequality. These variables were omitted from our analysis due to hardware and time limitations, so further exploration of the dataset is heavily encouraged.

Bibliography

Primary Sources

Dataset

"HM Land Registry: Price Paid Data", http://prod.publicdata.landregistry.gov.uk.s3-website-eu-west-1.amazonaws.com/pp-complete.csv (10/1/2019)

Secondary Sources

Data:

Nationwide Housing Price Index, 1973 - 2018, https://www.nationwide.co.uk/about/house-price-index/download-data#xtab:regional-quarterly-series-all-properties-data-available-from-1973-onwards

Halifax Housing Price Index, 1983 - 2018, https://www.halifax.co.uk/media-centre/house-price-index/

Retail Prices Index: Long run series: 1947 to 2019, ONS, https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/cdko/mm23

English Housing Survey: Housing stock report 2008, Department for Communities and Local Government

Text**:**

Atkinson A.B., “Unequal Shares”, Allen Lane The Penguin Press, 1972, 251

Clark G. & Cummins N., The big sort: The decline of northern England, 1780–2018, https://voxeu.org/article/decline-northern-england-1780-2018

Hamnett C., “Housing and Wealth in Britain”, Department for Communities and Local Government, Economics paper 6: Volume 1, 2010

Henley A., “Price Shocks, Windfall Gains and Hours of Work: British Evidence”, Oxford Bulletin of Economics and Statistics, 66 (4): 439-456

MacLennan D. & Tu, Y., “Changing Housing Wealth in the UK, 1985 to 1993: Household Patterns and Consequences”, Scottish Journal of Political Economy, 45 (4): 447-465

Levin E. & Price G., “Measuring Changes in Housing Wealth Inequality”, Department for Communities and Local Government, Economics paper 6: Volume 2, 2010

Piketty T., “Capital in the 21st Century”, Cambridge: Harvard University Press, 2014

Price G., “Has house price inequality risen or fallen since 1996?”, Department for Communities and Local Government, Economics paper 6: Volume 1, 2010

Roses J. & Wolf N., The Return of Regional Inequality: European Regions between 1900 and 2010, Economic History Working Papers

Saez E. & Zucman G., “Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data (No. W20625)”, National Bureau of Economic Research, 2014

Thomas B. & Dorling D., “Know Your Place Housing: Wealth and Inequality in Great Britain 1980-2003 and beyond”, Shelter

Travers T., Sims S., Bosetti N., “Housing Inequality in London”, Center for London, 2016: 24

“A Divided Britain? Inequality Within and Between the Regions”, The Equality Trust

Nationwide Housing Price Index, 2018, https://www.nationwide.co.uk/about/house-price-index/methodology#tab:Methodology

Appendix:

Table 1: Proportion of property types as a % of all properties

Type of Dwelling	Household Survey	Land Registry 2008	Land Registry Full
Terraced	28.6	30	30
Semi-Detached	26	27	27.8
Detached	26.8	20	23.2
Flat	18.6	23	19

Source: HM Land Registry, English Housing Survey: Housing stock report 2008