Testing for Housing Bubbles across London Housing Markets

Seth Harrison

A thesis submitted for the degree of

Doctor of Philosophy in Finance

Essex Business School

University of Essex

February 20, 2025

Abstract

This thesis comprises three papers in which I investigate the existence, nature, and effects of bubbles in London housing markets.

In the first paper, I applied a recursive unit root test following Phillips et al. (2011) and Phillips et al. (2015) (PSY) to four housing affordability ratios (price to rent and income ratios, and the user cost to rent ratio, calculated using the five year average and an optimised forecast for expected price growth) across London boroughs for the time period 2001 - 2019. With all of these, housing bubbles are identified around the Global Financial Crisis (GFC) period. With the first two ratios, but not the latter two, housing bubbles are identified in the post-GFC period in the mid to late 2010s. The key result of this is that the price to income and rent ratios can misidentify bubble periods due to them ignoring other fundamentals (i.e. interest rates).

In the second paper, I tested the rational bubble model across London housing markets using the coexplosive framework developed by Nielsen (2010) and Engsted and Nielsen (2012). I found that the data did not fit the rational bubble model for all London boroughs, which I argue is evidence against the existence of rational bubbles.

In the third paper, I tested for ripple effects across London boroughs using a spatio-temporal model following Holly et al. (2011). I found that there is evidence of price

convergence and dependence, however there is not evidence of Central London acting as a dominant region. I further tested the London housing markets with a coexplosive model of bubble transmission following Evripidou et al. (2022). I found that several regions showed signs of coexplosivity. I argue that this means the ripple effect could be explained by the spatio-temporal transmission of information.

Acknowledgments

Undertaking a PhD has been a remarkable journey, and I would not have been able to do it without significant support and guidance along the way. Firstly, I would like to thank Professor Simon Price, and my supervisors, Dr Sam Astill and Professor Neil Kellard, for their invaluable support. They have provided me with long hours of insightful discussion, guidance on tough theoretical problems, and indispensable academic and professional advice. Without them this journey would not have been possible. I am grateful to Dr Lazaros Symeonidis for his time and effort in chairing my supervisory panels. I want to extent my gratitude to my viva examiners, Professor Robert Taylor and Dr Emily Whitehouse, for their insightful comments and their thorough review of my thesis. I am sure that the quality of my thesis has been greatly improved by their expertise. I would also like to thank Professor David Harvey for his insightful discussion, in particular regarding the PSY procedure and bubble collapses.

I would like to thank my family for their unwavering support and endless patience throughout the years. I am particularly grateful to my brother Dr Judd Harrison for his insightful comments and proofreading of this thesis. Finally, I would like to thank my wife Wan-Yu Harrison, for her unending love and boundless support. Without her encouragement, I would never have embarked on this journey.

Contents

List of Figures

List of Tables

I. INTRODUCTION

Economic bubbles have existed for almost as long as financial markets have. Early examples of asset price bubbles include the tulip bubble which collapsed in 1637 and the South Sea bubble which collapsed in 1720. Asset price bubbles can be devastating for those who get caught up in them. Individuals can lose savings if they buy an asset in the peak of a bubble that subsequently collapses in value. In the case of housing, homebuyers can enter negative equity if they buy a house with a large loan and prices subsequently collapse. However the effects on the wider economy can be even more extreme. There is the wealth effect, where declining house prices decline household consumption (Case et al., 2005). Furthermore housing bubbles can also have wider macroeconomic effects - there is now general consensus that the 2008 Global Financial Crisis (GFC) was initially caused by a housing bubble in the United States (Baker (2008), Wheaton and Nechayev (2008) and Bernanke (2010) are some more prominent early papers in the literature discussing this). The collapse of the housing bubble, combined with the securitisation of mortgages, was a toxic combination that collapsed the entire world economy. Bubbles are one of the most pernicious economic phenomenon to exist. Yet despite this, the models and methods available to identify them are far from definitive. Furthermore there is not consensus on how bubble formation occurs. This thesis is structured around three papers which address issues in asset price bubbles, and housing bubbles in particular. These issues are primarily: the identification of bubbles, the existence of rational bubbles, and the spatio-temporal transmission of information causing price changes and bubbles to migrate geographically across housing markets.

The first paper of this thesis addresses the identification of housing bubbles. I discuss at length what a bubble is and the economics of housing markets to make clear how one might identify a bubble, and the current methods that are used in bubble identification. The key point is that a bubble is when prices increase in a self-fulfilling manner, because of expectations of future price increases, and moreover that these price increases are not justified by changing fundamentals. In the literature, most tests for bubbles revolve around testing for explosiveness in prices. The recursive right tailed unit root testing procedure developed by Phillips et al. (2011) and Phillips et al. (2015) is particularly popular. However any test for bubbles must take into account changing fundamentals - explosive prices by itself would not necessarily indicate a bubble. In financial assets, if a company announced particularly good, unexpected good news, its stock price might increase in an explosive manner. And this might well happen many times in a year - but this would not be a bubble because the price changes come from changing fundamentals. Likewise with housing, changing economic conditions can change house prices. It is perfectly possible for house prices to increase over many years, because of changing economic conditions (e.g. a shortage of housing being built), without there being a bubble. Because of this, it is common in the literature to see tests for explosiveness done not on price series, but some kind of affordability ratio. In housing, the two most common ratios this is done for are the price to income ratio, and price to rent ratio. However, as I discuss in the first paper of this thesis, this ignores many of the key drivers of house prices - in particular, interest rates. How exactly housing prices relate to fundamentals is what makes the identification of bubbles such a particularly difficult task. Even near the peak of the 2006 US housing bubble, there were many papers arguing that house price changes could be explained by changing economic fundamentals (in particular, interest rates) e.g. Himmelberg et al. (2005), Gallin (2008). While there have since been breakthroughs in some areas of bubble detection, these are largely econometric methods for detecting explosiveness in time series. This doesn't address the underlying issue: that explosive price increases can be explained by changing fundamentals. To address this, in the first paper of this thesis I suggest another affordability ratio that includes interest rates, house prices and rental prices: the user cost to rent ratio. The user cost of housing is the per period cost of owning a home and is a function of house prices and interest rates, among other things, that is commonly discussed in the housing economics literature (e.g. Duca et al. (2021), Hill and Syed (2016), Meen and Whitehead (2020), Miles (1994), Mulheirn (2019) *et cetera*). In theory, the user cost of housing should be equal to the services associated with housing (i.e. rent or imputed rent). I test all of these ratios at the regional level in London, for the time period 2001 to 2019. I show that by using the price to rent and price to income ratios, one would misidentify bubble periods where price increases can actually be explained by changing interest rates, in particular house price increases in the late 2010s, though with all the affordability ratios there were bubble periods identified before or around the Global Financial Crisis period.

The second paper in this thesis provides evidence regarding whether the bubbles are rational, or irrational. I examine in detail the rational bubble literature. A rational bubble is a bubble which takes place in a context with rational agents, and one in which rational agents would partake in. In a rational bubble, participants know that assets are overpriced, in the sense that prices are not commensurate to underlying fundamentals, yet it is nonetheless in their interest to take part in a bubble because expected returns on the bubble asset match their required rate of return. On the other hand, an irrational bubble is a bubble caused by agents acting in irrational manners - having excessive belief in their own ability, for instance. In this paper, I discuss various debates on whether a rational bubble *could* exist, even theoretically, and ultimately conclude that a rational bubble could theoretically exist, but in practice a more pertinent question is whether the bubbles which do actually exist are rational or irrational. In this paper I explicitly test the rational bubble model using the coexplosive framework developed by Nielsen (2010) and Engsted and Nielsen (2012). Again, I do this for London housing markets, at the regional level. This framework involves testing various restrictions that are implied by the rational bubble model, in a model that allows for explosiveness, with likelihood ratio tests. I find that the rational bubble model does not fit the data at all well. The key economic relationships that one would expect with the rational bubble model are comprehensively rejected. I argue that this is evidence against the existence of rational bubbles, and is evidence for the existence of *irrational* bubbles.

The third paper in this thesis addresses bubble contagion, where in this context a bubble transmits from one housing market to another, in association with the ripple effect. The ripple effect is a phenomenon in housing markets that is well documented in the literature, where changes in house prices in one region are subsequently followed by changes in house prices in neighbouring regions (Meen, 1996, 1999). This might be in price convergence - where prices in two regions share a long run relationship. So if one region's prices increase by a large amount, there will be a tendency for either neighbouring regions' prices to rise, or the region's prices to fall. There might also be a more general ripple effect in price dependence - where a change in one region is followed by a change in neighbouring regions, which might be in the same direction. If one region's prices increase, neighbouring prices might also increase and vice versa. Despite the ripple effect being well documented as an empirical phenomenon, there is still debate on what the causes and mechanisms of the ripple effect are. One apparent explanation is that it is caused by the spatio-temporal transmission of information - that good or bad news about the economy hits some regions first (and subsequently changes house prices), and this information radiates outwards geographically. There are other explanations however: for instance that what is actually changing in a spatio-temporal manner are the economic drivers of housing and the subsequent changes in house prices makes it appear that there is a ripple effect in house prices. For instance, incomes might change first in one region and then in surrounding regions in a ripple effect like manner, which then changes house prices. A similar story could be told for the effects of migration changing housing demand. In the third paper of this thesis, I estimate a spatio-temporal model for London boroughs, following Holly et al. (2011). I find evidence of price convergence and price dependence, though there is not evidence of Central London acting as a dominant region. A separate literature, though on a related topic, is developing on "co-explosivity", where a bubble will transmit across markets (Evripidou et al., 2022). The idea is that while there may be two bubbles across different assets and across time, the bubble in these assets is, econometrically, essentially the same underlying bubble. Thus in a manner similar to cointegration, the two asset prices will be linked in a fundamental relationship to each other. This would occur because the market participants in one market were developing expectations about price increases that lead to a bubble, and as information is transmitted in a spatio-temporal manner across networks, these expectations would then transmit to other markets where bubbles would form. Thus the mechanism of bubble contagion is most likely similar to one of the proposed mechanisms of the ripple effect: the spatio-temporal transmission of information. Therefore in the third paper of this thesis, I also test the coexplosivity model of Evripidou et al. (2022) and find evidence of coexplosivity across London housing markets at the regional level. I argue that this is evidence that the ripple effect occurs due to the spatial transmission of information.

II. THE LIMITATIONS OF UNIT ROOT TESTING PRICE AFFORDABILITY RATIOS

i. Introduction

AN economic bubble is where price increases in an asset are driven by expectations of future price increases. Prices are increasing, and this price increase is driven by pressure from speculators who are buying the asset because they expect the asset's price to go up in what becomes a self-fulfilling prophecy. The existence of housing bubbles can be significant - housing bubbles can have pernicious effects on both consumers and the wider economy, and it is widely thought that the 2008 Global Financial Crisis was ultimately caused by a housing bubble in the US[1].

House prices across the UK, and especially in London, have rapidly increased in recent years. It is common to read stories about the existence of a housing bubble and its impending crash. Figure 1 shows how house prices in London increased from 1995 to 2024, in real terms. From 1995 to 2008, house prices in London increased at a very high rate, before collapsing. This led many contemporaneous commentators to describe the UK housing market as having gone through a "bubble" (E.g. Bone and O'Reilly (2010), Clark et al. (2010), Hay (2009)). After this collapse, house prices again began rising sharply. By 2016, house prices had surpassed the 2008 level by a very wide margin, and they remained at this level for many years. Since 2022, London house prices have started to decline precipitously. Some commentators would posit that this confirms the existence of there being a bubble -

[1]Baker (2008), Wheaton and Nechayev (2008) and Bernanke (2010) are some examples of the many papers in the literature describing the pre-GFC housing bubble in the USA

there was a very sharp increase, which is perhaps starting to crash[2].

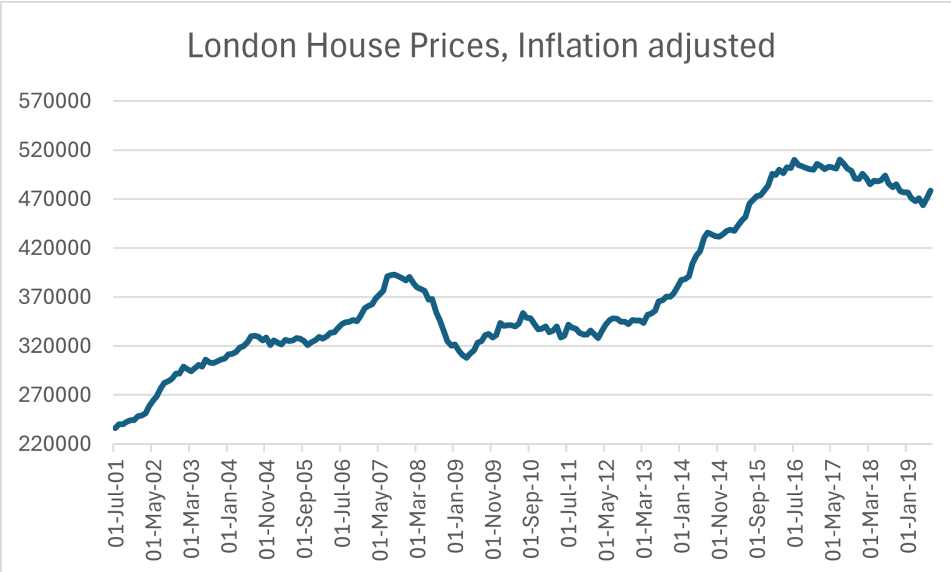


Figure 1: *London house prices*

However, just because house prices have increased does not mean that there is a bubble. Likewise, a crash in house prices does not mean that there was a bubble that has since burst[3]. A bubble is where prices are going up because of expectations about price increases. There may well be sound economic reasons that explain movements in house prices. One

[2]The housing market has been particularly volatile since the COVID crisis and the subsequent inflationary environment and monetary policy tightening. The housing market has faced numerous shocks over the last few years, such as real estate markets shutting down during COVID, house prices spiking when they reopened, and most recently higher mortgage rates. As I will discuss in the literature review, the housing market is characterised by frictions. Shocks to the housing market can take some time to become reflected in prices. Moreover, the volatility around the COVID lockdowns might be suggestive of structural breaks. For this reason the focus of this paper will be on the period of high house price increase up until the COVID crisis, i.e. until the end of 2019.

[3]Here and throughout this thesis, I consider a "crash" to be simply a large decline in prices, e.g. a 20% decline.

natural route is to compare prices to factors which affect affordability. Figure 2 shows house prices relative to income in London, the house price to income ratio. Clearly, house price growth has outpaced growth in income. Likewise, Figure 3 shows how London house prices have increased relative to rent. This shows how relative to renting a home, *owning* a home has become more expensive. Many commentators would find this sufficient to identify a housing bubble, and indeed, testing for bubbles via the price to income and price to rent ratios is ubiquitous in the literature, as I will discuss in the literature review.

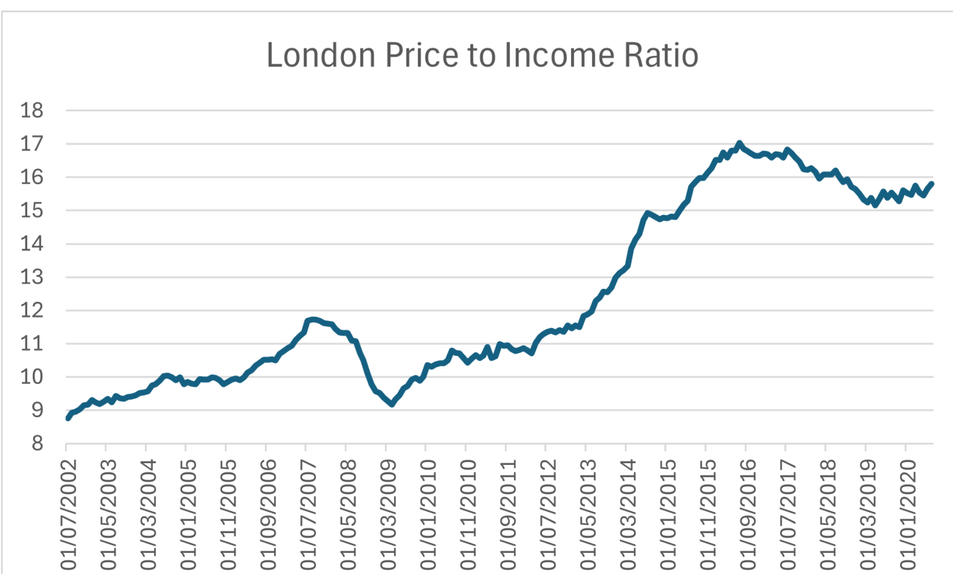


Figure 2: *London house price to income ratio*

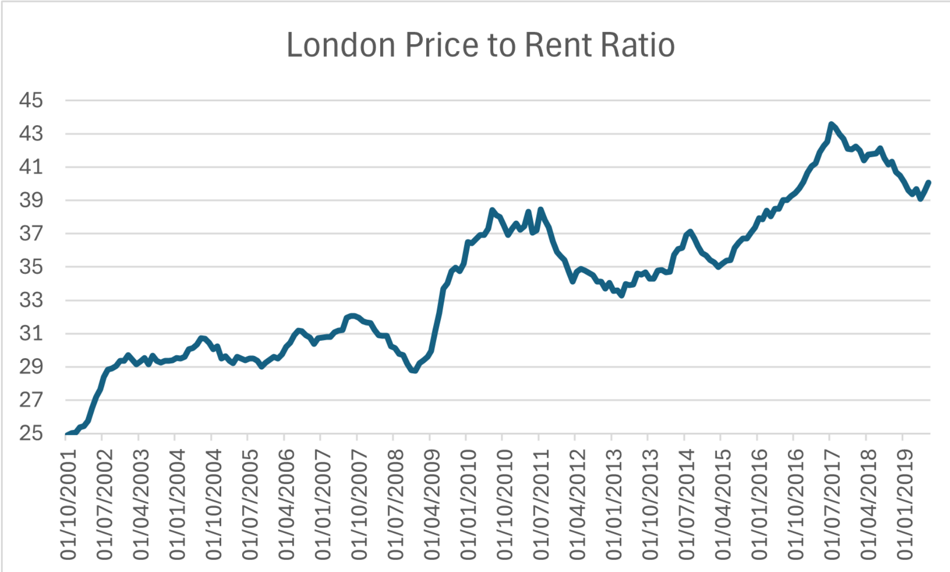


Figure 3: *London house price to rent ratio*

However, incomes and the price of rent are only a relatively small part of how housing prices are determined, and what it means for housing to be affordable. Most houses are purchased with a mortgage. The percentage of buyers who buy using a mortgage varies across the UK and across the world as a whole. In England & Wales, the percentage of homes purchased with a mortgage has almost always been above 60%. In London, it has typically been above 70%. Even in regions of London like Kensington & Chelsea, where house prices are very high, and houses are more likely to be left vacant[4] (which would make it a likely area where there could be a housing bubble), around half of houses are typically purchased with a mortgage. Figure 4 shows how the percentage of homes purchased with a mortgage has varied over time since 2012, in England & Wales, London, and Kensington &

[4]According to the most recent census data in 2021, 19% of houses in Kensington & Chelsea are truly vacant (i.e. vacant, and not used as a second home). https://www.ons.gov.uk/peoplepopulationandcommunity/housing/bulletins/numberofvacantand secondhomesenglandandwales/census2021 accessed 12/06/2024

Chelsea.

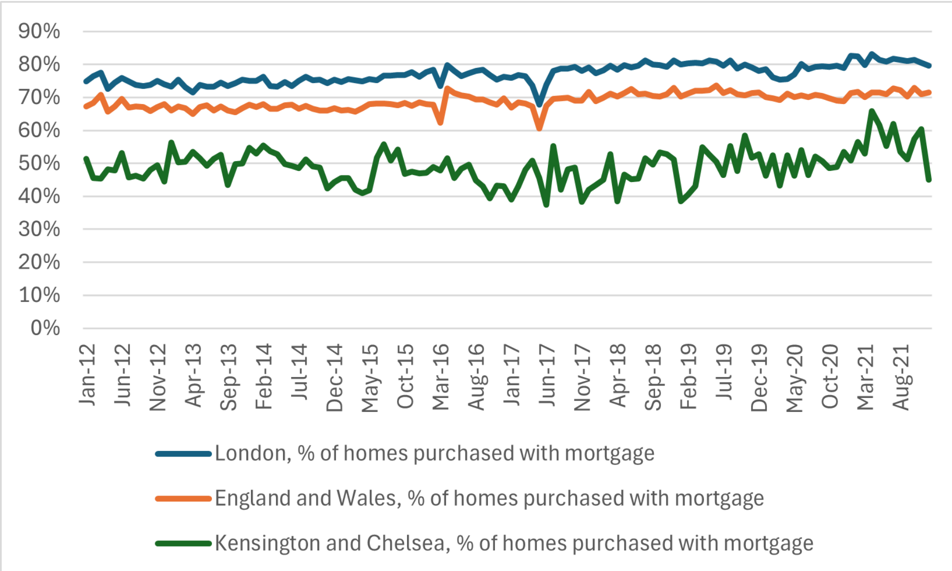


Figure 4: *Housing transactions across various UK regions, with % financed by mortgage*

Over the last twenty years, mortgage rates (and interest rates more generally) had been decreasing (Figure 5 shows how UK interest rates have changed over time). Prices are inversely related to interest rates. As mortgage rates go down, home buyers who purchase with a mortgage can afford a greater loan amount, all things being equal, which will allow them to bid up the nominal price of a home for the same monthly mortgage payment. Moreover, even with cash purchases, the same dynamics that affect mortgage rates will change the discount rate of a cash buyer, changing prices in much the same way[5].

[5]For instance, consider a standard discounted cash flow model: $P_t = E_t[\sum_{i=1}^{\infty}(\frac{D_{t+i}}{1+R_{t+i}})]$, where P_t is the current price, D_t are the cash flows associated with the asset (e.g. dividends), and R_t is the discount rate (the required rate of return on the asset). In most asset pricing models the discount rate is equal to the risk free rate plus some risk premium. If the risk free rate (i.e. the interest rate set by the central bank) is changed, this directly affects the discount rate. For instance, suppose the interest rate goes down so that the discount rate also goes down. This means that future cash flows are discounted less, and are thus worth more in the present. This means that the present value of discounted future cash flows is higher, and thus the price is

So London house prices have increased massively, even relative to income and rents - yet over the same period, up until very recently (late 2022), mortgage rates and interest rates more generally had also gone down. Clearly to answer the question of whether there was a housing bubble requires an in depth discussion of the nature of the housing market and the economic drivers of how housing prices are determined. This is what I review in the Literature review of this chapter, in Section ii. I also discuss econometric methods to detect housing bubbles. In the rest of this paper, I test for housing bubbles in London at the borough level. To do this, I applied the recursive unit root test procedure developed by Phillips et al. (2011) and Phillips et al. (2015) to several housing affordability ratios: the price to income ratio, price to rent ratio, and user cost to rent ratio. Section iii describes the recursive unit root methodology in detail, and how the affordability ratios were calculated. Section describes the data that I used. In Section v I describe the results. Section vi concludes, with some remarks about areas for future research.

higher. The inverse would happen if interest rates increased - as the discount rate goes up, future cash flows are discounted more and are worth less in present value terms, which makes the price go down. *Alternatively*, another way of thinking about this is in practice with what happens to asset's returns as interest rates change. If interest rates were very high, e.g. the yield on a risk free government bond was 6%, then investors would require a higher return that this to purchase a risky asset. They would not be willing to purchase a risky asset if its expected return was less than 6% plus an appropriate risk premium. They would simply invest in government bonds instead. This would force the asset's price down, until it did meet the required expected return. Likewise, if government bond yields are very low, this will mean investors are willing to pay higher prices to get returns elsewhere. If the expected return on a risky asset were much higher than the risk free rate plus risk premium (which would be the case if prices were low), this would make the risky asset an attractive investment. As investors flock to buy the risky asset, this will force its price up, until the expected return is equal to the risk free rate plus an appropriate risk premium.

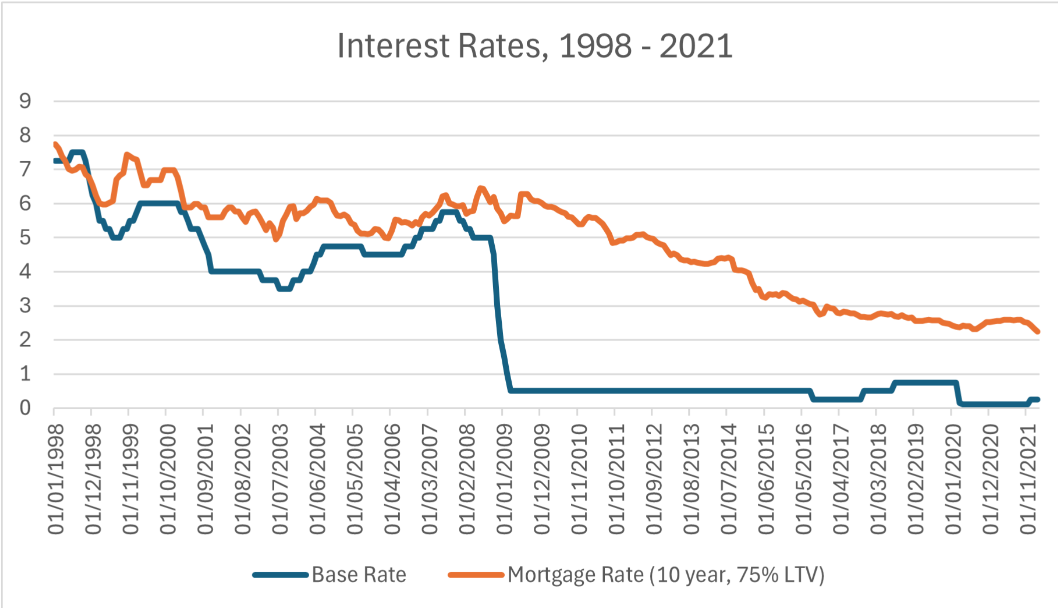


Figure 5: *UK interest rates over time*

ii. Literature Review

What are the Fundamental Drivers of Housing Prices?

As I mentioned in the introduction - a housing bubble is not just where housing prices have gone up or become unaffordable (though obviously this is likely to happen during a positive bubble). It is where housing prices have gone up because speculators expect prices to keep going up and hence a profit can be made by buying housing (and this buying pressure increases prices further), and because of this housing becomes removed from its fundamental determinants[6]. But what are the fundamental determinants of the housing market, and how are prices determined by them?

[6]This definition of a bubble is paraphrased somewhat from Stiglitz (1990)

The Nature of the Housing Market The first thing to note about how housing prices are determined is the nature of the housing market - compared to other markets, it is characterised by large frictions - transaction costs, informational asymmetries and taxes (Duca et al., 2021; Miles, 1994). When buying a home, a person will either require a large loan, which they will spend much of their life paying off, or they will require a large sum of cash, which may represent most of their life savings. Often there are large chains, where a buyer must wait until they have found a buyer for their own home, and that buyer must also find a buyer for their home, et cetera. There are transaction costs associated with estate agents, taxes (i.e. stamp duty in the UK), and with moving furniture and other belongings. Furthermore, there is a high level of heterogeneity in housing markets. No two houses are the same. One of the main drivers of a house's price is the area it is in: an area with local amenities like schools and good transport links will have higher house prices than a house in an area without these. Likewise, housing is more expensive in some countries and cities compared to others. Apart from area, one of the main determinants of the value of an individual house is its features – such as number of bedrooms, bathrooms, condition of electricity and plumbing etc. Even the house's style can affect its value, and often the features of an individual house can be totally unique. Finding the right house, and finding the right buyer for a house, can take time. There are also potential informational asymmetries, as a house buyer will likely have less information about a house than the person selling it. Homebuyers will typically pay to have a survey done on the property they purchase, but again this raises further questions about finding a competent surveyor etc. So in addition to transaction costs, there are also search and matching costs in housing markets.

Because of these frictions, changes in the fundamentals underlying the market can take some time to be reflected in house prices.

Furthermore there is good evidence that housing markets are highly predictable. For instance, transaction volume can be used to predict changes in house prices, i.e. a fall in transaction volumes will be followed by a fall in prices. For instance Tsai (2018) uses a VECM model to show that under most circumstances transaction volume has significant information for future house price appreciation in the UK housing market. There is evidence that shocks in fundamentals will produce large transitory shocks in transactions, and simultaneously small but persistent shocks in prices. This has been found in the UK (Andrew and Meen, 2003), US (Berkovec and Jr, 1996) and Dutch housing markets (de Wit et al., 2013), for instance. One explanation for this is that there is asymmetric information about the demand side of the market, but not the supply side: when an individual unit of housing is listed on the property market, typically the seller or an estate agent will base the house price on what similar properties in the same area recently sold for. Likewise a related idea is that house sellers are loss averse, and will be less willing to sell houses for a lower price than they bought for. So if there is a change in fundamentals, so that demand is lower, the house will still be put on the market for the same price as before, so it will take longer to be sold. As houses stay on the market for longer, the number of transactions is lower. This can be used to predict the fall in house prices that eventually comes from the change in fundamentals (de Wit et al., 2013; Meen and Whitehead, 2020). This is not totally undisputed however, as some research has found a negative relationship between house prices and transaction volume, such as (Arslan et al., 2015). There is some evidence that there are predictable price trends in countries, where a city will act as a ‘lead’ city, and will increase (or decrease) in price, followed by an increase (or decrease) in price in other cities. For instance, in the UK, there is evidence that London acts as a lead city for most of the rest of the country, with areas of the country closer to London being affected more strongly (Holly et al., 2011; Meen and Whitehead, 2020).

There is even evidence that house prices can be predicted from historic data alone for both the UK and the US, and that house prices in these countries display autocorrelation. (Andrew and Meen, 2003; Schindler, 2013, 2014; Tsai, 2018).

The key takeaway from this is that we can expect many of the traditional tests for market inefficiency - i.e. predictability- to come back as positive when tested on housing markets. Some argue that predictability in housing markets is not an indication of market inefficiency. For instance, Miles (1994) sets out a model with forward looking rational agents maximizing utility, subject to budget constraints, and shows that the model predicts supply and demand shocks to housing to increase/decrease prices in a predictable way as prices and supply adjusts to a new equilibrium, and Miles (1994) argues that because of this, predictable price changes in housing are not proof that housing markets are inefficient.

Income, Demographics & The Available Housing Stock I discussed briefly how the price of an individual house, taken individually, is determined by its location and features. For instance, a real estate agent would value a house being newly brought onto the market by comparing it to what similar properties sold for, and this would be based on location and features. On a larger scale, such as across different times or different countries, location becomes much more important in determining housing prices. This is because at this level housing prices are explained by the economic conditions of the local housing market. One of the primary determinants of housing prices is the area's aggregate income, given the available housing stock. All things being equal, if the population of an area increases, then housing prices will increase. Likewise, if average incomes increase, then house prices increase.[7] A good example of the effect of income on housing prices are the high housing prices in Silicon Valley. Housing prices in Silicon Valley are extremely high, however incomes are also extremely high as well. For instance, in the San Jose-Sunnyvale-Santa Clara Metro area, 2019, the median value of owner-occupied housing was $1,116,400, compared to the US national average of $240,500[8]. Obviously the house prices in Silicon Valley are much higher than average. However during the same period and in the same area median household income was $130,865, which was almost double the national average of $65,712[9]. The high housing prices in Silicon Valley are at least partly explained by the high incomes in Silicon Valley.

Changes in income and the available housing stock can come from multiple sources. The most obvious are the general economic conditions affecting people's ability to pay for housing. However, many changes in income and housing stock availability have historically come from unexpected sources, which are not directly associated with housing (Meen and Whitehead, 2020, p.4-5). For instance, post-World War Two there were advances in transportation technology which have made a daily commute into city centres from suburbs possible. This allowed the creation of the suburbs in the USA post-World War Two which massively increased the availability of housing and increased the population that

[7] Also important is household financial wealth, increases in which will increase house prices independently to changes in income. (Meen and Whitehead, 2020, p.43-44)

[8] U.S. Census Bureau (2019). Value American Community Survey 1-year estimates. Retrieved from <https://censusreporter.org>

[9] U.S. Census Bureau (2019). Household Income in the Past 12 Months (In 2019 Inflation-adjusted Dollars) American Community Survey 1-year estimates. Retrieved from <https://censusreporter.org>

was willing to live in the area that was formerly the suburbs. This had a significant effect on house prices. For instance, in 1950 the median home price in the United States was \$7,354[10], compared to an average family income of \$3,300[11]. This gives a price to income ratio of 2.2 - which would be considered extremely low by modern standards. Conversely, financial deregulation in the 1980s had the eventual side-effect of increasing availability of mortgage financing which, as I will explain later, contributed to rises in housing prices.

Demographic changes can also affect demand for housing, as aggregate income is not determined by average income alone but also total population. This concept is particularly popular in non-academic media[12], who speculate on ways that changing demographics can affect housing prices. For instance, if a country has an ageing population, then perhaps prices might adjust to better reflect the preferences of old-age pensioners and holiday homes might appreciate ahead of the average house. There is some evidence for these kind of trends in the literature. For instance, Mankiw and Weil (1989) find that the introduction of the baby-boomer generation into adulthood and the housing market in the 1970s in the US caused a large increase in housing prices, though this result is controversial. More recently, in the UK, Levin et al. (2009) find that the growth rates of the 20-29 and 30-44 demographics have a significant effect on house prices, and use this to explain differences in house price growth between England and Scotland. There are many other demographic factors that are relevant for housing prices. For instance, in the UK, household size has

[10] U.S. Census Bureau accessed via <https://www.census.gov/data/tables/time-series/dec/coh-values.html>

[11] U.S. Census Bureau accessed via <https://www.census.gov/library/publications/1952/demo/p60-009.htmll>

[12] For instance, https://www.investopedia.com/articles/mortages-real-estate/11/factors-affecting-real-estate-market.asp

roughly halved in size over the last century (Holmans, 2005). This has largely been caused by a greater proportion of households being formed by single people, and this has increased housing demand directly by increasing new household formation. (Barker, 2004).

As I stated earlier, one of the main determinants of housing prices is aggregate income given available housing stock. In theory, if the available stock of housing increases fast enough, then this can offset any increases in income, increases in population, or changes in mortgage affordability. Most government reports on housing affordability in the UK have focused on the apparent lack of housing stock given the UK's population, and have focused on how government policy can increase the levels of home-building by increasing planning permission, and the speed with which granted planning permission results in new housing being built (Barker (2004) and Department of Communities and Government (2017), for instance). In practice however, it is not usually possible to increase the rate of construction of newly built houses to completely off-set house price increases, particularly in countries such as the UK (which Barker (2004) fully acknowledges). There is some truth to the government's claims that house prices have increased because not enough houses have been built - levels of construction of social housing have decreased since the 1970s (Meen and Whitehead, 2020). But most studies have found that the amount of new housing that would have been required to maintain house prices would have to have been extremely high, and probably intolerable given the environmental and ecological impact (Barker, 2004).

The main factor which determines levels of construction for new housing is the availability of land, in particular land that can be built on (Meen and Whitehead, 2020). Many countries have restrictive regulations controlling the construction of new properties, and the proportion of new build to existing properties is usually very small. Countries with stricter building regulations experience greater price increases over time, under certain circumstances. This is because areas with high regulations tend to have lower price elasticity of supply. In these areas, which includes the UK, when housing prices start to increase, housing supply will not increase to match the price increases. Prices will just go up (Glaeser et al., 2008; Meen and Whitehead, 2020; Duca et al., 2021). In the UK there are heavy regulations about building on greenfield land, and because of this, new builds will typically increase the total housing stock by only 1% each year. Because of this, there are some in the literature, for instance Mulheirn (2019), who argue that the UK's housing issues simply cannot be solved with the construction of new housing. Even if construction of new-build housing was to double in the UK, the yearly effect on the total housing stock would be very small. According to calculations by Mulheirn (2019), if the UK built 300,000 net additions a year (which it has never done) for 20 years, house prices would decrease by at most 13%, which is obviously tiny compared with actual house price appreciation.

Credit Availability & Government Policy The other primary determinant of housing prices is the availability of credit. Most new homeowners pay for their home via mortgages[13]. If the mortgage expense that a mortgage borrower needs to pay in a given amount of time goes down, given a nominal house price, what will usually happen is that the nominal house price will go up until the mortgage expense has matched its previous value. For instance, if interest rates decrease, then the mortgage rate will also decrease (mortgage rates follow interest rates closely). If the mortgage rate is lower, then given the same

[13] According to the English Housing Survey 2020-2021, 95% of first time buyers purchased their home with a mortgage. However since 2013-2014, most households are actually outright owners. This is mainly isolated to older demographics however - 63% of outright owners are retired.

nominal house price, monthly payments will be lower. This allows the house buyer to bid up the nominal price of the house, until the monthly mortgage payments match their old value (Meen and Whitehead, 2020; Duca et al., 2021).

There are many other factors that affect the availability of credit as well as interest rates. For instance, the way the mortgage market operates. In the UK before the 1980s, the mortgage market was dominated by building societies. There was a shortage of mortgage funding available, so mortgage lending was constrained (Meen and Whitehead, 2020). This is included theoretically as a 'shadow price' in the user cost of housing (Dougherty and Order, 1982; Meen, 1990). Borrowers could not borrow large mortgages even if they wanted to. In the 1980s there was a period of financial liberalisation, and mortgages became much easier to obtain. Borrowers could now choose to take out a large mortgage if they so wished. As a result of this the stock of mortgage debt relative to household income has more than quadrupled since 1970 in the UK[14]. This can also be seen by how mortgage sizes compare to the required down-payment, and income of the mortgage buyer. Before the Global Financial Crisis, for instance, around 4% of mortgages were above 95% Loan to Value, and more than 3.5 times yearly income. These types of mortgages were extremely rare outside of this period, such as since the GFC or before the financial liberalisation of the 1980s (Meen and Whitehead, 2020). So as a result of mortgage market liberalisation, the total amount that people could get in mortgage financing, given their income level, went up. Hence people could afford to pay higher nominal housing prices, which allowed housing prices to be bid up. High total mortgage debt is correlated with high housing prices, however causation is not clear. It is probably the case that both increased total mortgage

[14] According to data from the ONS and Bank of England, as cited by (Meen and Whitehead, 2020)

debt and increased housing prices are caused by increased availability of credit (Meen and Whitehead, 2020).

In fact, anything that lowers the actual cost of buying a mortgage will allow mortgage buyers to bid up nominal prices. Government policies and subsidies can also affect the ability of mortgage-buyers to pay their debts. For instance, the government's Help-to-Buy scheme. This reduces the financial burden on first time buyers, which in turn increases their ability to pay higher nominal prices for homes. Help-to-Buy has been described as like "throwing petrol onto a bonfire"[15], in terms of its effect on housing prices. Likewise, the same applies to any other government grants, policies and subsidies that reduce the financial burden on mortgage buyers. If their financial burden is reduced, they can afford to pay a higher nominal price for housing, which raises prices. An example of this is that during the COVID-19 pandemic, there were a number of measures introduced to maintain housing while people were furloughed or quarantined i.e. foreclosure and eviction moratoriums. These, combined with increased demand for living space and other government policies that maintained incomes during the pandemic (such as direct transfers), have led to increases in house prices in most developed countries since the start of the pandemic (Duca et al., 2021). There is also evidence that housing benefits can affect housing prices. Government housing benefit subsidises rental payments for lower income households, and there is evidence from the UK, USA and France (Gibbons and Manning, 2006; Fack, 2006; Susin, 2002) that much of this will be captured by increased rental prices for let properties. Through the relationship between rental prices and house prices, there is evidence from the UK that increases in rental payments from increased housing benefit can increase housing prices

[15] Sam Bowman, Executive Director of the Adam Smith Institute, October 2017

(Braakman and McDonald, 2020).

There has been some work that suggests that it is not only how much the mortgage actually costs that affects house prices, but also the mortgage buyer's estimation of how much the mortgage will cost them (Sustek, 2021). Mortgage buyers effectively borrow from future income, so future expected income is important in mortgage financing (Meen and Whitehead, 2020). For instance, if the mortgage buyer has expectations that their income will increase, then the perceived cost of the mortgage will be lower, and hence they will be willing to bid up housing prices by getting a larger mortgage.

The Relationship between Housing Prices & Rental Prices The user cost of housing is the actual cost of owning a unit of housing in a period. It can be shown under several assumptions (such as if the purchase is financed through a mortgage) with a rational agent consumption optimisation framework (Miles, 1994) that the user cost of housing is given by:

$$UC_t = P_t(r_{m,t} + \delta_t + \tau_{p,t} - E(\dot{P}_t)) \qquad (1)$$

Where P_t is the price of a unit of housing, $r_{m,t}$ is the mortgage rate, δ_t is the depreciation rate (i.e. maintenance costs), $\tau_{p,t}$ is property taxes associated with owning a home. $E(\dot{P}_t))$ is the expected capital gains from owning the unit of housing. The terms in the brackets are also sometimes known as the user cost of capital for housing (given that they are a real interest rate)[16].

[16] Note, that this can be derived with different sources of capital contributing to the house purchase (i.e. if the house purchase is partly financed with cash, as most home purchases are). The cost of capital for cash used to finance the purchase is the opportunity cost of not investing in an alternative asset e.g. the stock market. Call

Housing prices and rental prices are theoretically linked, by equating the user cost of owning a home with the stream of rents at the rate that the homeowner would rent out the home to themselves at (Meen and Whitehead, 2020; Poterba, 1992). This is also known as imputed rent. The imputed rent is the service that owning a house provides - one does not have to pay rent because of home-ownership. In practice it is not possible to observe imputed rent directly however, so often rental market data is used instead (Meese and Wallace (1994) for instance). There are several justifications for this – first, if renting is significantly cheaper than buying a home, then people who want somewhere to live will opt for renting a home, and if purchasing a home is relatively cheap, then more people will buy homes instead.

The second justification is from the perspective of investors – if the rental yield on properties was particularly low, then investors who buy homes to let will switch to other investments. This ought to lower the price of housing as investors sell housing, and increase the price of rent as fewer homes are available to rent. But if rent is particularly high relative to prices then investors will tend to buy houses more, which will increase house prices and decrease rental prices. An implication of this is that prices ought to be equal to rent discounted by an appropriate discount rate, as with other financial assets. If this discount rate is linked to mortgage rates, then this present value relationship can be made equivalent

this cost of capital $r_{e,t}$, and the user cost of housing becomes $UC_t = P_t(\rho r_{e,t} + (1-\rho)r_{m,t} + \delta_t + \tau_{p,t} - E(\dot{P}_t))$, where ρ is the proportion of the house purchase financed with cash. For simplicity, I will be assuming that $\rho = 0$, or, equivalently, $r_{e,t} = r_{m,t}$. Now this assumption might seem unrealistic - however strictly speaking the opportunity cost would be in an investment with the same level of risk, and in theory two assets with the same level of risk ought to have the same expected return. So this later assumption is arguably not too unrealistic.

to equating the user cost of housing to rent. There are some criticisms of this, however. If someone was to rent out their property to themselves, they would probably have a different price compared to the rental market due to informational asymmetries on the latter. A tenant not paying rent is a major concern for landlords and estate agents, and a premium is given to trustworthy tenants. There is evidence that landlords will not raise rent on existing tenants as often as they do for new tenants – because keeping an existing trustworthy tenant is valuable. Likewise, renters using a house face restrictions to their behaviour that owner-occupiers do not (Meen and Whitehead, 2020).

Furthermore, there is a difference of quality between homes occupied by owner-occupiers and homes occupied by renters. Rental properties tend to be smaller, and are more likely to be flats, while owner-occupier homes tend to be larger and higher quality. Therefore, data issues could potentially arise from using rental market data.

Housing Bubbles & their Detection

As I have discussed previously, a housing bubble is not just when housing prices are high. It is when they are so high that they have become removed from the fundamentals. I have discussed the fundamental factors that determine the price of housing - primarily income given the available housing stock, the availability of mortgaging financing, and the level of rents compared to the level of housing prices. Most housing bubble detection tests in the literature involve testing if housing prices have become removed from the fundamentals in a bubble-like fashion. However, there is much variation on both what fundamentals and what relationship are tested, and how the relationship is tested.

Some housing bubble detection tests focus on the time-series properties of housing prices, as it is often thought that a bubble will have explosive price behaviour. For instance it has been shown by Diba and Grossman (1987) that with a rational framework and constant discount rates, prices will have an explosive component. Because of this some bubble detection tests are on price alone, suggesting that explosive house price increases might be evidence of a housing bubble (For example Zhi et al. (2019) and Zhou and Sornette (2006)).

Obviously however sometimes rapid price increases can be justified by changes in fundamentals. It is common therefore to test for bubbles by testing the time series properties of an economic variable related to the price and fundamentals, which we would expect to be non-explosive or stationary. If this variable is non-explosive or stationary, then there are no bubbles. This could then be tested with an Augmented Dickey Fuller (ADF) test. An early example of this is Meese and Wallace (1994). Meese and Wallace (1994) impose a present value relationship on house prices, where the price of a house is the discounted stream of future rents, discounted by the user cost of capital for housing. They then test the forecast error implied by the model for stationarity with an ADF test. (Other examples of such a present value relation being used for housing in the literature include Engsted et al. (2016) and Mikhed and Zemčík (2009)). One criticism of this however is an ADF test will fail to detect a bubble that has collapsed during the time period being tested, or multiple bubbles which have periodically built up and collapsed (Evans, 1991).

In response to this, a recursive unit root testing procedure was developed by Phillips et al. (2011) and Phillips et al. (2015). This repeatedly applies ADF tests recursively to the data and can time-stamp periods of explosive bubble-like behaviour. It is often referred to as the PSY procedure. However it is still very important that the series that is tested would be expected to be non-explosive. Most previous attempts at applying the PSY testing procedure for detecting housing bubbles have been on a variable which excludes important variables, which could account for a rapid change in prices without a bubble, such as the price to rent ratio or the price to income ratio.

For instance: Engsted et al. (2016) test the price-rent ratio of OECD countries, including the UK. They conclude that their results indicate explosive behaviour in the UK from 2001 - 2005. Rherrad et al. (2021) test price to rent ratios in several metropolitan areas in Canada and conclude that there is evidence of bubbles in several markets. Tsai and Chiang (2019) test the price to rent ratio to several cities across China, and conclude that there is evidence of bubbles, that starts first in the top tier cities and then spreads to second tier cities. Huang and Shen (2017) test the price to rent ratio in Hong Kong and conclude that there is evidence of a bubble. Greenaway-McGrevy and Phillips (2016) test both the price to rent and price to income ratios in New Zealand at the regional level and They conclude that there is evidence of bubbles. Petris et al. (2022) test London Housing boroughs, similar to this current paper, but testing price and price to income. They conclude that there is explosive behaviour in Barking and Dagenham, Newham and Bexley. Hu and Oxley (2018) test the house price to income ratios at the state level in the USA. They conclude that there is evidence of there being bubbles in several states. Yip et al. (2017) test the price to CPI ratio in Malaysia (arguing that imputed rent is a component of CPI) and they identify periods as having bubbles.

Testing the price to rent alone ratio ignores interest rates - if interest rates go down, so do mortgage rates, and if mortgage rates go down then for the same monthly payment then mortgagors can afford a higher nominal house price. Hence lowering interest rates increases house prices, and interest rates have been decreasing consistently over the last twenty years. So ignoring interest rates ignores an important fundamental which could lead to rising house prices without there being a bubble. Likewise this is true with the price to income ratio. Furthermore with the price to income ratio, it is possible that the price of owning a home is simply becoming more expensive. Widespread home ownership is only a relatively recent phenomenon.

In this paper I test the user cost of housing to rent ratio. As I explained before, these are theoretically linked, and furthermore this accounts for most of the important variables that determine house prices i.e. interest rates and rental prices. This procedure makes no mention of incomes or housing stock. However, following, Mulheirn (2019) I would argue that higher incomes or less availability of housing will be reflected in the rental prices: a distinction can be made between the cost of housing services (i.e. the shelter, comfort etc. that housing provides) and the actual cost of housing as an asset. Rental prices measure the former, while house prices measure the latter. If house prices were high due to a genuine lack of housing, this would also be reflected in the rental prices (Mulheirn, 2019).

There are of course many other ways in the literature to test for housing prices which I have not discussed yet here. For instance, some housing bubble detection tests test for a cointegrating relationship between house prices and supply and demand variables (i.e. a linear combination of these variables and prices is stationary). Often these factors will be relatively wide ranging compared to those used in other methods. Examples include population, income and construction costs (e.g. Clark and Coggin (2011)). If these supply and demand factors are cointegrated with house prices, this is interpreted as the prices being proportionate with the fundamentals. If not, then a housing bubble cannot be rejected. There has been some criticism to this approach in the literature (e.g. Gallin (2006)), given that a cointegrating relationship ought to exist only if a number of assumptions about the relationship between house prices and these factors hold, such as the price elasticity of supply and demand variables being stable (which may not be the case).

I also mentioned earlier the present value relation used by Meese and Wallace (1994) and others. This present value approach treats housing more like traditional financial assets and is thus complimented by a corresponding literature in empirical finance of testing present value models for traditional financial assets. There are many ways of testing this present value model (both for housing and financial assets). One popular method is by testing a cointegrating relationship implied by the model, by fitting a present value model into a vector autoregression framework. This was originally proposed by Campbell and Shiller (1987, 1988a) who focus on the spread, defined as $S_t = P_t - (1/R) * D_t$ (where for stocks, P_t is price, R is the required return, assumed to be constant, and D_t is dividends at time t). The efficient market hypothesis implies this would be a cointegrating relation. This VAR methodology has been expanded upon by Nielsen (2010) and Engsted and Nielsen (2012) to include coexplosive processes in addition to the cointegrating process. This method allows for the present value model to expanded to include an additional term for the possibility of rational bubbles. Basically, similar to how a standard VECM model will break down differences into a cointegrating vector and lagged differences, the coexplosive cointegrating VAR model breaks down the differences of differences from the explosive process into a cointegrating and coexplosive vector, and lagged differences. This coexplosive VAR methodology has been applied to housing markets (i.e. Engsted et al. (2016) and Kivedal (2013)). However, Engsted et al. (2016) follows the method used by Engsted and Nielsen (2012) very closely – they assume that the required rate of return, R, is constant, and estimate this parameter from the model. This might be appropriate with a present value model for the stock market, but for the housing market, the required rate of return is highly dependent on interest rates. Indeed, lowering interest rates are usually cited as the main factor for rising house prices. So assuming the required rate of return to be constant leaves out one of the most important factors for prices. Kivedal (2013) does address this issue, by constructing a series for rent that is rent deflated by interest rate. However this still leaves out many other factors that are relevant for the owner occupier cost of housing, such as expected capital gains, depreciation etc. Moreover, even adding these terms, simply deflating the rental series doesn't accurately capture the relationship between rent and prices. $R_t/((1+i_t+\alpha))$ is obviously not equal to $R_t/((1+i_t)(1+\alpha))$, which is what this implies. They find that house prices are not explained by a present value model, even with a bubble.

Additionally, another way of testing the present value relation, more generally, is by testing the predictability of future returns and rent growth on the rent-price ratio. In the empirical finance literature, it has been shown by Campbell and Shiller (1988b) that if the present value model and the efficient market hypothesis hold (and there are no bubbles), then the dividend-price ratio can be log-linearised into the sum of future expected returns and expected dividend growth. This is interpreted as the dividend-price ratio summarizing expectations for future dividend growth and/or returns and implies that the dividend-price ratio should be a good predictor of one or both. This has been applied with the same present value logic to the housing market, with rents replacing dividends. For instance, Engsted and Pedersen (2015) apply this method to OECD countries, including the UK, and Gallin (2008) and Cochrane (2011) applies this to the United States. Traditional wisdom suggests that the dividend yield (and rent-price ratio) should be an indicator of future dividends (or rent). For instance, if a stock has a low dividend yield, the reason investors are willing to pay the prices they are is because they expect future high dividends. However, in practice most empirical results are that the dividend yield/rent-price is actually a good predictor of future returns. So, when the rent-price ratio is low, this is usually followed by low price increases, and vice versa. It could be argued this is evidence that house prices go through distinct periods of under- and over-pricing. Some of these other methods will form the basis for later chapters in my overall thesis.

iii. Methodology

The PSY test procedure

The PSY test procedures revolve around recursively testing periods of the data with right tailed unit root tests. The unit root tests are *right tailed* because we are testing if the series is explosive, as opposed to stationary. For instance, consider the standard ADF test equation:

$$\Delta y_t = \delta y_{t-1} + \sum_{i=1}^{k} \alpha_i \Delta y_{t-i} + \varepsilon_t \tag{2}$$

If the process y_t is unit root, then δ should be equal to zero, and less than zero if it is stationary. The standard ADF is thus a left tailed test - if δ is significantly less than zero, then the null hypothesis of y_t being unit root is rejected in favour of y_t being stationary. Here however we are considering a variable that could be explosive - in which case δ is *greater* than zero. Hence with the right tailed ADF test, if δ significantly *greater* than zero, then the null hypothesis of y_t being unit root is rejected in favour of y_t being *explosive*.

As the whole sample period is unlikely to be explosive in the case of a bubble, only the period where the bubble forms, the PSY test procedure applies ADF tests recursively to different testing periods of the original sample. How exactly this is done depends on the exact test.

With the Backwards Supremum ADF test procedure (BSADF) the test statistic is calculated as follows: denote the end of the period being tested as time r_2. The start of the sample is denoted r_1, and the ADF statistic calculated from r_1 to r_2 is denoted $ADF_{r_1}^{r_2}$. With r_2 fixed, the ADF test statistic is calculated repeatedly for different values of r_1, first for r_2 minus a minimum window length[17], denoted r_0. So the first ADF statistic calculated is $ADF_{r_2-r_0}^{r_2}$. The window length is incremented by one, and the ADF test statistic calculated again. ADF test statistics are calculated repeatedly in this manner, until r_1 is equal to the start of the series. The BSADF statistic is the supremum (maximum value) of all these ADF statistics i.e.

$$BSADF_{r_2}(r_0) = \sup_{r_1 \in [1, r_2 - r_0]} ADF_{r_1}^{r_2} \tag{3}$$

The BSADF statistic is calculated repeatedly, for every possible value of r_2 (from the minimum window length until time T). We can then datestamp which periods had bubble like behaviour by identifying where the BSADF exceeds its critical value. Phillips et al. (2015) show that the estimator for the datestamping is consistent.

[17]The minimum window length is $\delta Log(T)$. This requirement is included to prevent short lived blips from affecting the results of the test. The δ term is included to adjust for data frequency

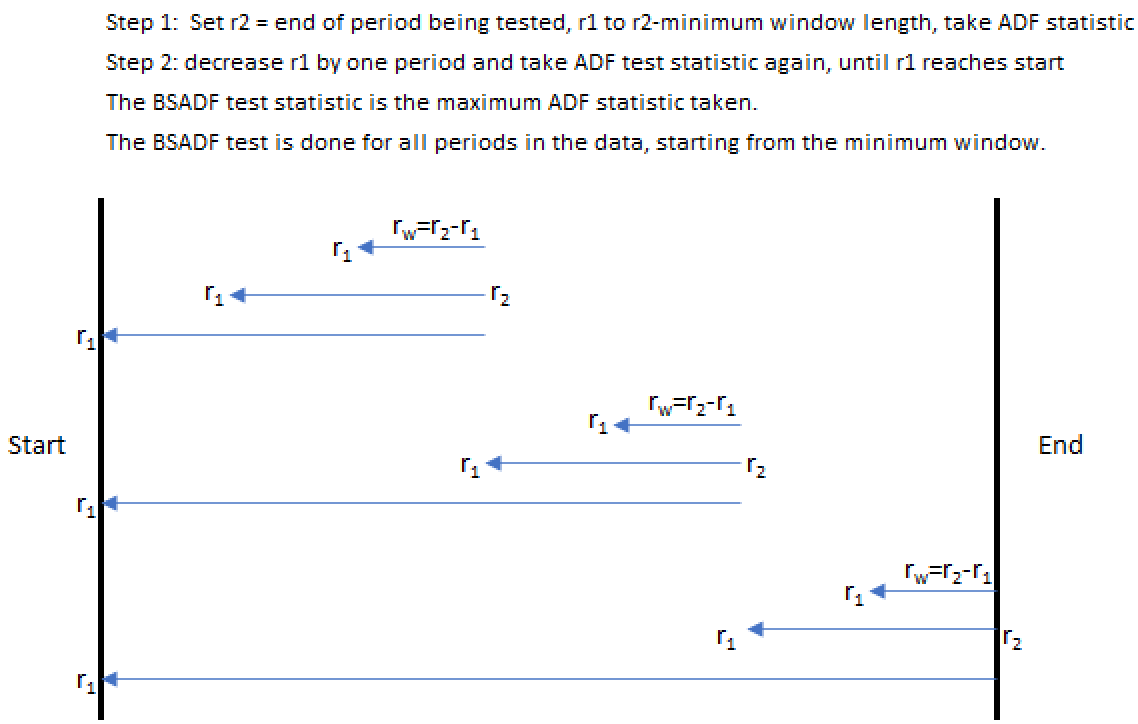


Figure 6: *An illustration of the BSADF test procedure*

The procedure for the Generalised Supremum ADF (GSADF) test is similar. The start of the test period, r_1, is set to time 1, and the end of the test period, r_2, is set to r_1 plus a minimum window length, r_0, and the ADF test statistic is calculated. r_2 is incremented by one and the ADF test statistic is calculated again. This is repeated until r_2 is equal to the end of the whole sample, T. This whole procedure is repeated again with r_1 at time 2, then time 3 etc. until r_1 is equal to time $T - r_0$. The GSADF test statistic is the supremum of these ADF test statistics:

$$GSADF_{r_2}(r_0) = \sup_{r_1 \in [1, r_2 - r_0], r_2 \in [r_0, T]} ADF_{r_1}^{r_2} \tag{4}$$

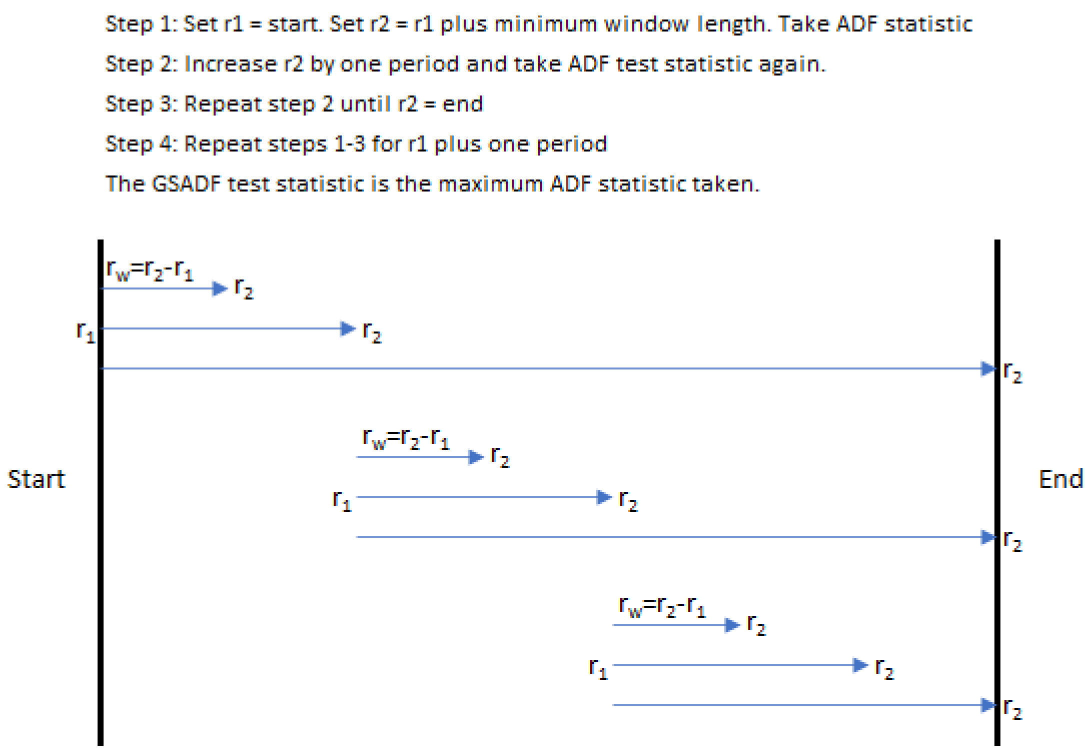


Figure 7: *An illustration of the GSADF test procedure.*

These procedures are illustrated with Figures 6 and 7, inspired by a similar figure from (Phillips and Shi, 2020, pg.7).

The Wild Bootstrap Procedure

Harvey et al. (2016) have shown that the critical values of the original PSY test can suffer from size distortions in the presence of heteroskedasticity. Because of this, critical values for the tests were generated using a wild bootstrap process following Phillips and Shi (2020). The wild bootstrap procedure goes as follows:

Step 1: First we estimate the following model over the full sample:

$$\Delta y_t = \mu + \sum_{i=1}^{p} \phi_i \Delta y_{t-i} + e_t \tag{5}$$

This is a essentially an ADF equation like Equation 106, but estimated with the restriction that $\delta = 0$.

Step 2: Let τ_0 be the minimum window length, and τ_b be the number of observations that the size is to be controlled for. For a sample size of $\tau_0 + \tau_b - 1$, a bootstrap series is generated by:

$$\Delta y_t^b = \sum_{i=1}^{p} \hat{\phi}_i \Delta y_{t-i}^b + e_t^b \tag{6}$$

The initial values of y_t^b for $t = 1, ..., p+1$ are generated as $y_t^b = y_t$. The bootstrap residuals are generated by $e_t^b = w_t e_l$, where w_t is standard normally distributed $w_t\ N(0,1)$, and e_t were estimated in the ADF-style regression in Step 1. This addresses any heteroskedasticity that might be present in the data. Likewise the parameters $\hat{\phi}$ are the OLS estimates from Step 1.

Step 3: The series of PSY statistics are computed from the bootstrap series, $\{PSY_t^b\}_{t=\tau_0}^{\tau_0+\tau_b-1}$. The maximum value of this series is stored:

$$M_t^b = \max_{t \in [\tau_0, \tau_0+\tau_b-1]} PSY_t^b \tag{7}$$

Step 4: Steps 2-3 are repeated for $B = 499$ bootstrapped series.

Step 5: The α significance level critical value of the PSY methodology is thus the $1-\alpha$ quantile of the values $\{M_t^b\}_{b=1}^{B}$.

Producing Figures

To highlight periods of explosive, bubble like behaviour, I used the BSADF function.

Periods of explosive, bubble like behaviour are highlighted by where the BSADF statistic exceeds the critical value which was obtained in the wild bootstrap process. Formally this can be written as an indicator function as follows:

$$I(r_2) = \begin{cases} 1 \text{ if } r_2 \in r_2 : BSADF_{r_2}(r_0) > CV_{r_2}(\alpha) \\ \\ 0 \text{ otherwise} \end{cases} \tag{8}$$

It is also worth noting that the BSADF procedure can also highlight periods of price *collapse* as explosive as well. See Appendix vii for an explanation of why this occurs. Obviously any asset bubble is associated with explosive price *increases*. This is why we are testing for explosiveness. If a series is explosive in a technical sense, but prices collapse over this period, this is ignored and not highlighted as a bubble period.

Affordability ratios

To find evidence of explosiveness in house prices relative to their fundamentals, I test several affordability ratios using the PSY tests. I first test the price to rent ratios, and price to income ratios, as these are commonly used in the literature. As I have explained in the literature review however, these ignore changes in interest rates (among other things) that are crucial in determining house prices. So I also test an affordability ratio that accounts for all of these factors, i.e. the user cost of housing to rent ratio. As discussed in the literature review, the user cost of housing is the per time period cost of owning a house. It is equal to the price of the house times the cost of capital. The cost of capital contains many factors i.e.

the interest (mortgage) rate, depreciation, council tax, transaction costs, risk premium, and expected capital gains on housing. I do this for 22 different London regions (boroughs and boroughs merged together) and London as a whole.

The user cost to rent ratio can be written as follows:

$$y_{i,t} = UC_{i,t} / R_{i,t} \tag{9}$$

Where UC_i, t is the user cost and R_i, t is the imputed rent, for each region i at time t.

$$UC_{i,t} = P_{i,t} \left(r_{m,t} - E(\dot{P}_{i,t}) \right) + \delta_t + \tau_{p,i,t} + \pi_t T_t \tag{10}$$

Note that the maintenance costs, δ_t, and mortgage rate r_m, t do not vary across regions. Usually the property tax is expressed as a rate, multiplied by property price, but here I include it separately as the data is based on the average amount paid per property. Likewise with depreciation and transaction costs. Where the price variable was an index, I scaled this up to be proportionate to housing prices at the start of the period. For the price to rent and price to income ratios, the income and rent is yearly, to make the analysis most intuitive.

Constructing the User Cost of Housing

The user cost of housing contains several variables which are not directly observed, i.e. the risk premium associated with owning housing, and the expected capital gains on housing. These need to be calculated to construct the user cost of housing. Furthermore the treatment of transaction costs requires some consideration: transaction costs with housing are high, so it is unlikely for a homeowner to own a house for one month and then immediately sell it. So adding the total transaction cost to the user cost each period would result in a user cost which is unrealistically high.

Transaction Costs Instead, I calculated the per period transaction cost by treating moving home as a hazard which has a probability of occurring each period, and if it does occur it imposes some cost. For instance without transaction costs, house prices ought to obey:

$$P_t = E_t[\frac{s_{t+1} + P_{t+1}(1 - \delta_t - \tau_{p,t})}{1 + r_{m,t} + \alpha_t}]) \tag{11}$$

This is because the price of an asset should be equal to the sum of expected services that flow from owning it. If one were to own a home for one period, and then sell it at the end of the period (obviously an unrealistic assumption if the period was short), this would mean that the price would be equal to the expected housing service one receives from home ownership next period, s_{t+1}, plus the price of the house next period, P_{t+1},deflated by depreciation, δ_t, and taxes $\tau_{p,t}$. The required return is the relevant interest rate, $r_{m,t}$ plus a risk premium α_t, if there is one.

By rearranging this for the flow of services from owning housing, s_{t+1}, and equating the flow of services from owning housing to the user cost, we get the equation for the user cost of housing:

$$UC_t = P_t(r_{m,t} + \delta_t + \tau_{p,t} + \alpha_t - E(\dot{P}_t)) \tag{12}$$

For more details on the derivation, see Appendix vii

With transaction costs, as a per period hazard which imposes a cost, the condition becomes

$$P_t = E_t[\frac{s_{t+1} + P_{t+1}(1 - \delta_t - \tau_{p,t}) - \pi_t T_t}{1 + r_{m,t} + \alpha_t}]) \tag{13}$$

Where π_t is the probability of a transaction occurring, and T_t is the cost of a transaction. The user cost then becomes

$$P_t(r_{m,t} + \delta_t + \tau_{p,t} + \alpha_t - E(\dot{P}_t)) + \pi_t T_t \tag{14}$$

So with transaction costs, the user cost increases each period by simply the probability that a transaction occurs in a period, multiplied by the cost of a transaction. The cost of a transaction is assumed to have two components - stamp duty, which is calculated based off of property price (the rates and property bands vary year to year), and legal and surveyor fees which are assumed to be 1% of property value. The probability that a transaction takes place is assumed to be the property turnover for that year divided by the dwelling stock.

Risk Premium One commonly cited paper in the literature which calculates the user cost of housing is Himmelberg et al. (2005). For the risk premium, they cite Flavin and Yamashita (2002) as having estimated the risk premium of owning housing at 2.0%. These two papers have in turn been cited elsewhere in the literature when the user cost of housing is being estimated, e.g. Hill and Syed (2016), Chen et al. (2022), Kulikauskas (2017), Finicelli (2007), Dolinar (2018) etc. Unfortunately it is not clear where Himmelberg et al. (2005) get the figure of 2.0% from. Flavin and Yamashita (2002) analyse the portfolio optimisation problem of a typical household - who can invest in financial assets (i.e. stocks and bonds), and owner-occupied housing. They take the proportion invested in owner-occupied housing, which is taken as given[18], and demonstrate how this will affect the proportions of the different financial assets in the household's portfolio. They do not

[18]Based on real data of how different age groups tend to have different proportions of their wealth invested in housing.

calculate a risk premium for housing[19].

I attempted to calculate the risk premium using several approaches. The first is the Consumption CAPM formula. There are several versions of this in the literature. For instance Bailey (2005) has the risk premium equal to

$$\frac{\gamma cov(r_j, c)}{1 - \gamma E(c)}$$

where r_j is the return on risky asset j, and c is the growth rate of consumption, and γ is the coefficient of relative risk aversion. However Campbell et al. (1998) have the risk premium equal to:

$$\gamma cov(r_j, c) - var(r_j)/2$$

These are close to each other however: $1 - \gamma E(c) \approx 1$, and $var(r_j)/2$ may be small compared to $\gamma cov(r_j, c)$, so $\frac{\gamma cov(r_j,c)}{1-\gamma E(c)} \approx \gamma cov(r_j, c) \approx \gamma cov(r_j, c) - var(r_j)/2$. I decided to use the Campbell et al. (1998) version because it is computationally simpler. I assume γ to be equal to 2. Initially, I computed this over a rolling 5 year period. The resulting risk premium is very small, about 0.1%, and often actually *negative*. This is much lower than the level found in the literature, which is around 1-2% (Hill and Syed, 2016; Himmelberg et al., 2005). I tried with yearly returns instead, however the result is much the same. I also computed a CAPM risk premium too using the FTSE All-Share Total Return index, for comparison, and this resulted in a similarly small, sometimes negative risk premium, (because the covariance between housing and the market is often small and negative). The root cause of this is that the covariance/correlation between housing returns and the factors

[19] Though various other statistics are calculated which are relevant, such as the covariance matrix of housing and various financial assets. Likewise there is some discussion on the value of the coefficient of relative risk aversion, which is obviously not the same thing as the risk premium, but can be used in calculating it.

that are most commonly used in calculating risk premiums, i.e. consumption growth or market returns, is very close to zero (a result which has been found in the literature, i.e. Flavin and Yamashita (2002)). This is a somewhat counter-intuitive result, however it has been argued that the risk premium to owning housing should not be too high, as owning housing provides a hedge against rental prices (Sinai and Souleles, 2005). I ultimately decided to exclude the risk premium from calculations for the main results because of this. For robustness, I repeated the analysis for the user cost to rent ratio (calculated with five year average expected price growth), with the risk premium calculated in various ways: i) Fixed at 2% (which as previously discussed is common in the literature) ii) Fixed at 5% iii) Using a CAPM risk premium iv) Using a Consumption CAPM risk premium. The results of this are presented in Appendix vii. These different specifications result in no substantive changes to the results.

Expected Price Growth From the user cost equation, expected house price growth enters negatively. The intuition for this is that it is a benefit to hold an asset that is expected to appreciate in value. Hence expected capital gains on housing reduces the cost of home ownership. This creates a problem however in constructing the user cost of housing (that other authors have also remarked upon, for instance (Hill and Syed, 2016)), in that the user cost is incredibly susceptible to how the expected price growth is calculated. If one bases it off of a short sample length of past data, the user cost can often go negative. This is because housing prices often go through periods of rapid price increases. Depending on how the expected price growth is calculated, it can be so large that it outweighs the rest of the user cost. And because it enters negatively, then the user cost can become negative. From the point of view of economic theory, this seems absurd. Owning housing is having an asset that provides a stream of services. This should come with an associated cost. Thus the sample length of data used to calculate the expected price growth needs to be long enough so that short term bursts in growth do not make the expected price growth too high.

There is also a question of how to calculate the expected price growth. One option would be to take the simple average over a rolling window of past data. Another option would be to construct an optimal price growth forecast with a time series model, based on available data at the time. This is what I did initially - first I estimated ARMA(p,q) models of the form:

$$y_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{q} \theta_i u_{t-i} + u_t \tag{15}$$

with varying values for p and q, between zero and two, using the data over the period February 1995 to March 1998. I constructed forecasts from these models i.e.

$$y^f_{t+1} = \hat{\mu} + \sum_{i=0}^{p} \hat{\phi}_i y_{t-i} + \sum_{i=0}^{q} \hat{\theta}_i \hat{u}_{t-i} \tag{16}$$

I then chose the model specification which had the forecast with the lowest Root Mean Square Error (RMSE) over the period April 1998 to June 2001. RMSE is defined as:

$$\sqrt{\frac{1}{T}\sum_{t=1}^{T}(y^f_{t+1} - y_{t+1})^2} \tag{17}$$

Then I estimated the model again over the period February 1995 to June 2001. I did this for each region and used each model to construct forecasts for house price increases. Doing this for a one month ahead forecast resulted in the expected price growth rapidly fluctuating between positive and negative (which seems absurd). An implication of this is that with the expected price growth calculated in this way, the user cost is also highly volatile, and it rapidly fluctuates between positive and negative (which seems especially absurd). An example of this is given in Figure 8.

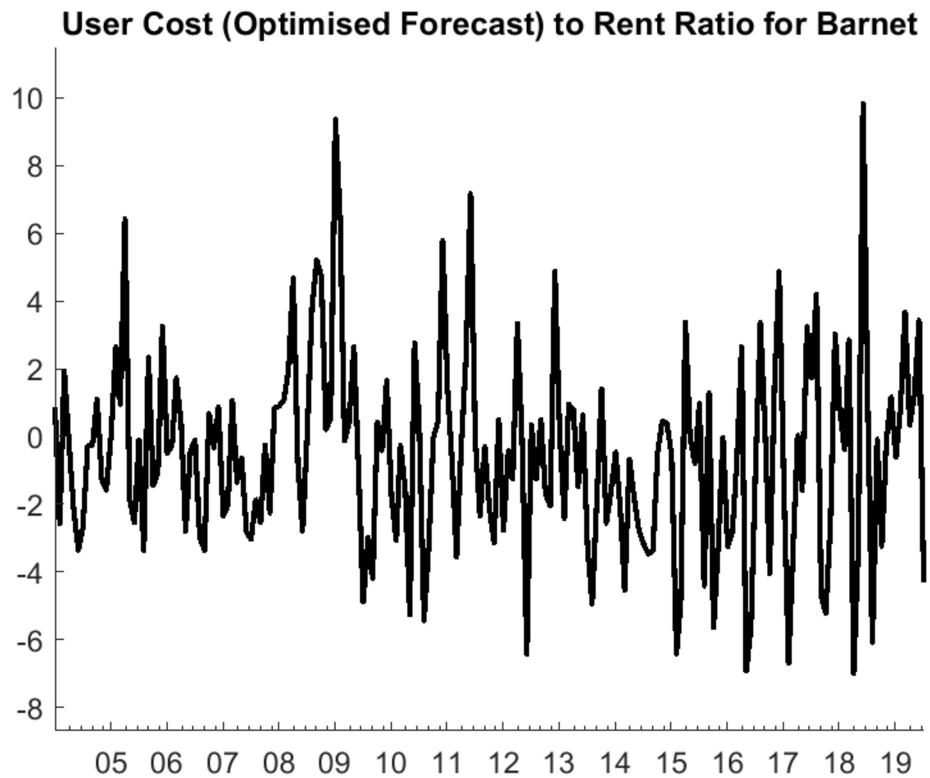


Figure 8: *User cost (optimised (one-month ahead) forecast) to rent ratio for Barnet*

Perhaps part of the problem is that homeowners do not buy and sell a home within one month, given how high transaction costs are. So instead I tried forecasting the expected price growth over one year ahead (and adjusted the data frequency accordingly), given this is a more reasonable length for a homeowner to occupy a house for and then sell. This also accounts for any seasonality left in the data.

So I estimated the one year ahead price growth forecasts with ARMA(p,q) models based on which forecast has the lowest RMSE, as before:

$$y_{t+12}^{12,f} = \hat{\mu} + \sum_{i=0}^{p} \hat{\phi}_i y_{t-i}^{12} + \sum_{i=0}^{q} \hat{\theta}_i \hat{u}_{t-i} \tag{18}$$

Note the lags in the ARMA(p,q) model this second time. Here y_{t+12}^{12} is the yearly growth rate, i.e. the growth over the next year. We are estimating growth twelve months ahead, using information known at the present. Hence I am estimating y^{12}_{t+12} with data known at time t or lagged from time t. I used data over the period February 1996 to March 1998 to estimate the models. I then chose the model specification which had the lowest mean square error over the period April 1998 to June 2001. I estimated the models again over the period February 1996 to June 2001 and used this to construct a forecast which I used as the expected price growth.

It is debatable whether it is correct to use an optimised forecast, or a simple moving average for the expected price growth. One could argue that because the expected price growth is the expectation of a typical representative agent making forward looking decisions, it should be a simple average. A typical consumer does not have access to the econometric models to make these calculations, and it seems hard to believe that they could have this knowledge implicitly. For completeness I calculated the user cost with both an optimised forecast and a simple average forming the basis of the expected price growth. As I discussed earlier, the length of time that one uses to calculate the expected price growth is crucial - so when doing the simple moving average, I calculated it using the last 5 years of data. This is the longest time-frame that can be done at this level, because borough level data only starts at 1996.

iv. Data

The UK House Price Index is produced by the Office for National Statistics (ONS) and for England uses sales data from HM Land Registry. It is a quality adjusted measure, that accounts for different houses being sold with different features. The HPI is just an index, and not a price series, so some of the figures would look unrealistic if the HPI was used in isolation (for instance the price to rent and income ratios). So where relevant, I scaled up the HPI to house price level by using the value of the Average Sales Price (for all properties), which is also produced as part of the HPI, on 01/07/2001. The HPI also includes statistics on sales volume per borough, and sales volume by transaction type (i.e. if the purchaser is buying with cash or mortgage).

Data for rental prices came from the imputed rental data produced as part of the Gross Value Added statistics published by the ONS. This is calculated using private rental market data produced by the Valuation Office Agency (VOA) and dwelling number figures produced by the Department for Levelling Up, Housing and Communities (DLUHC, formerly the DCLG), and accounts for differences across region and dwelling types. As the ONS produced figure is an aggregate across regions, I used data on dwelling stock over time produced by DLUHC to calculate imputed rent per dwelling.

The imputed rental data is merged together for some of the boroughs (e.g. Bexley and Greenwich are together despite being two boroughs). I merged the other data series together using either sales figures or dwelling stock, depending on how the figure per housing unit was calculated. The HPI is based on sales volume, while council tax is based off dwelling stock.

Data for interest rates came from the Bank of England. For the mortgage rate I initially used the series IUMBV42. This is the longest fixed length mortgage (5 years), at an LTV of 75%, where data is available that covers the entire sample. Initially I also used the Bank of England Base Rate to compare and contrast the effect from a different interest rate, however I found the results unrealistic so I do not include them. The 5 year 75% LTV rate is quite low - getting to less than 2% at the end of the sample. This seems too low for a representative mortgage rate for the aggregate market. The reason this rate is so low is because 5 years is quite a short period for a mortgage, and in reality mortgage periods tend to be much longer, switching to a variable rate after a period of fixed rate. The Bank of England does produce a series for a 75% LTV 10 year mortgage. There are however gaps in this series. For gaps of one month I linearly interpolated the data. For longer gaps I estimated what the series would have been using the two, three and five year mortgage rates. I first did ADF tests on all four, and it seems they are likely unit root. So I did a regression:

$$\Delta 10\_YEAR_t = \alpha + \beta_1 \Delta 2\_YEAR_t + \beta_2 \Delta 3\_YEAR_t + \beta_3 \Delta 5\_YEAR_t + \varepsilon_t \tag{19}$$

I then used this to estimate values for the 10 year 75% LTV rate when values were missing.

Data for income came from the ONS, from data produced as part of the Annual Survey of Hours and Earnings (ASHE). This is available yearly, so the data was linearly interpolated to be monthly. The income data for the City of London is very poor, so this is excluded from the analysis for data using income. The population data is only available from July 2002, so this is where series using incomes start.

Stamp Duty rates and bands are published by HMRC. I made some assumptions when calculating stamp duty: the purchase is not in a disadvantaged area (though this doesn't change bands too much), the purchase is for a freehold or leasehold property with more than 21 years remaining on lease, and the property is neither from a first time buyer nor for a second home, and it is not being purchased by a corporate body. I calculated the stamp duty cost based on the scaled up HPI, for each month and each borough.

For consumption, which was used in estimating the CCAPM risk premium, I used Household final consumption expenditure data, which is published by the ONS as part of the national accounts figures.

The ONS produces data on average council tax paid per dwelling. This enters the user cost directly, instead of as a rate, as mentioned in the methodology.

For depreciation, as part of the Consumer Price Index (CPI), the ONS produces an index for "regular maintenance and repair of the dwelling" as a part of "owner occupiers' housing costs". An actual figure was obtained for this amount from the Living Cost and Food (LCF) Survey, which could then be used to produce an actual number for the monthly depreciation cost, for the whole series. Of note is that the figure produced is very low - around £400 per year per dwelling. Perhaps this is unrealistic - for instance most tabloids list a maintenance figure of at least 1% per year[20]. For a house, which are typically priced in the hundreds of thousands of pounds, this would mean a yearly depreciation cost of thousands of pounds a year. By comparison the £400 figure seems absurdly low. Of course it is possible that this £400 figure is actually correct, as it only includes the direct maintenance costs associated with home ownership only. It excludes things such as cleaning supplies, and furniture maintenance for instance.

Finally, all price and cost data was deflated by CPI, which is produced by the ONS, to be in real terms. All price series are in levels.

v. Results

Price to Income Ratio

The first affordability ratio which I tested was the price to income ratio. The bubble periods identified via BSADF for London are shown in Figure 9. Clearly the ratio shows a general upward trend. The first bubble period that is highlighted is around the 2008 Financial Crisis,

[20] e.g. https://www.thebalance.com/home-maintenance-budget-453820

specifically in 2007. After this, there is a clear drop in the price to income ratio (as house prices dropped over this period, faster than incomes). The ratio reaches a low in 2009, and after this starts to rise again. It gets to very high levels in 2014, and another bubble period is highlighted that lasts into 2017. Now it is worth noting that with using the price to income ratio, the period post-2017 is not highlighted as being a bubble, even the level of the price to income ratio is still much higher than the period before 2014. A bubble is fundamentally a period of overpricing. However if house prices were overpriced in 2014, according to the price to income ratio, why are they not still overpriced in 2019? Moreover, what is the level at which housing become overpriced? There is no *a priori* suggestion with the price to income ratio. Figures 10-12 show the identified bubbles using the price to income ratio at the borough level. The price to income ratios have the same general trend as London in all the boroughs - i.e. a rise around 2007, followed by a collapse around 2009, followed by a much higher rise over the next ten years. Some of the boroughs do not show as clear a trend as others. The rises and drops in Camden are generally less clear, for instance. Some of the boroughs did not get identified as going through a bubble around the 2008 crisis, though most did. The boroughs that were *not* identified were Bexley, Camden, Greenwich, Hackney, Hillingdon, Hounslow, Sutton, Tower Hamlets, and Westminster. Every single borough (apart from Camden) had a bubble period identified in the mid-2010s period. GSADF test results for all regions are reported in Table 1[21], and mirror the results with BSADF.

[21]Critical values are again generated following Phillips and Shi (2020).

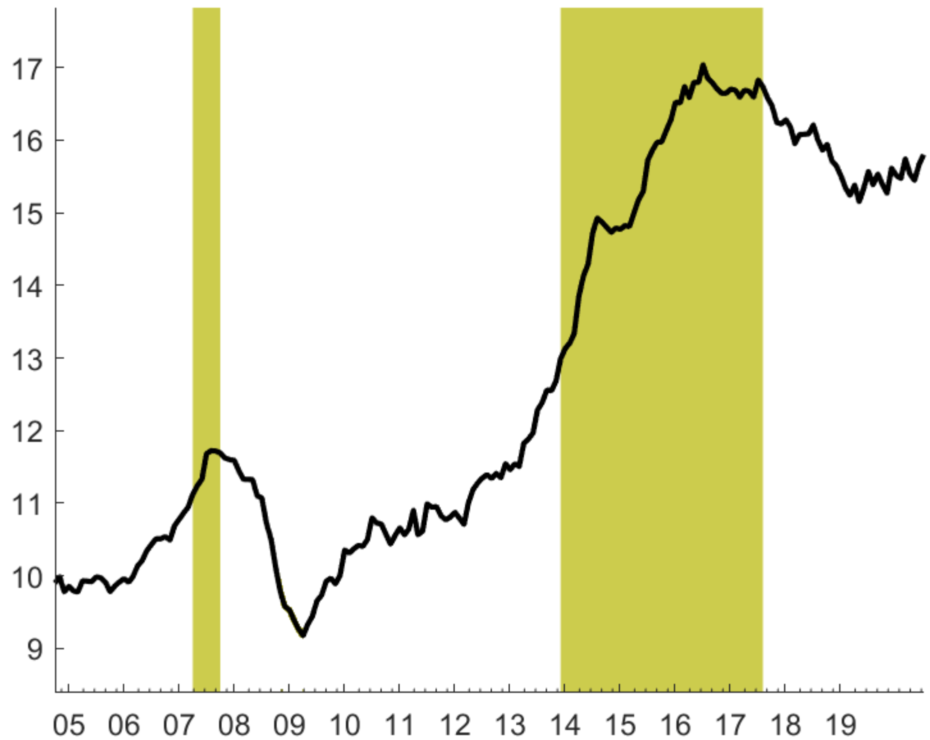
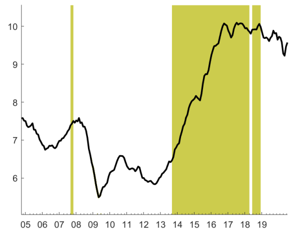
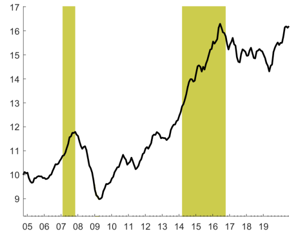


Figure 9: *Price to income ratio for London*

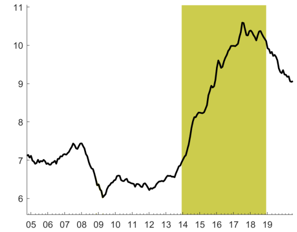
The next affordability ratio that I tested was the price to rent ratio. Figure 13 shows the price to rent ratio for London. It shows the same general upward trend as the price to income ratio, though with some differences. There is a small rise and fall between 2007 and 2009, however the first substantive rise and fall occurs in 2009 and ends around 2013, before rising again to a very high level, peaking in 2017, before collapsing post 2017. Again, there are two bubble periods highlighted. The first bubble period identified starts in 2009, and peaks around 2010, before collapsing in 2011. So like the price to income ratio, the price to rent ratio indicates a housing bubble occurring during the 2008 Financial Crisis, though the exact timings differ. It also identifies a bubble during the mid-2010 period, with similar timings (i.e. a peak in 2017).



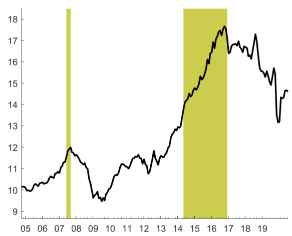
(a) *Price to income ratio for Barking and Dagenham*



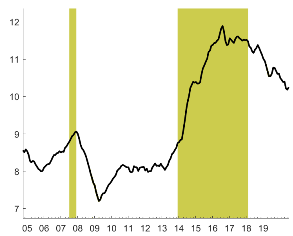
(b) *Price to income ratio for Barnet*



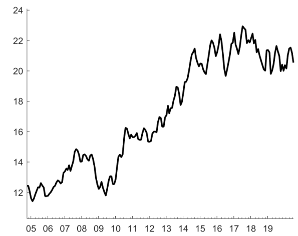
(c) *Price to income ratio for Bexley*



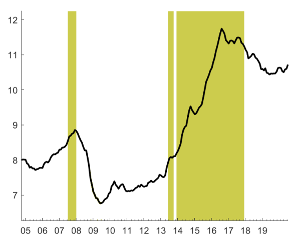
(d) *Price to income ratio for Brent*



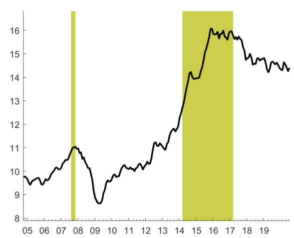
(e) *Price to income ratio for Bromley*



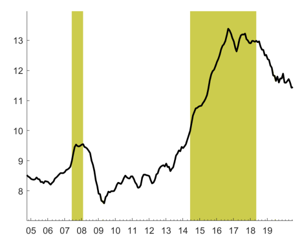
(f) *Price to income ratio for Camden*



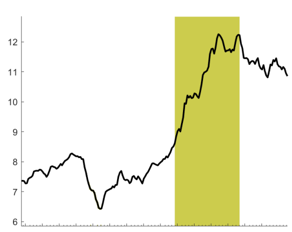
(g) *Price to income ratio for Croydon*



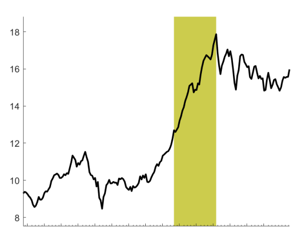
(h) *Price to income ratio for Ealing*



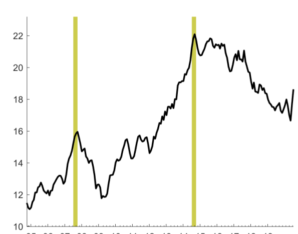
(i) *Price to income ratio for Enfield*



(j) *Price to income ratio for Greenwich*

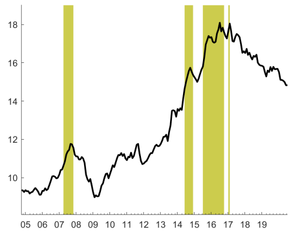


(k) *Price to income ratio for Hackney*

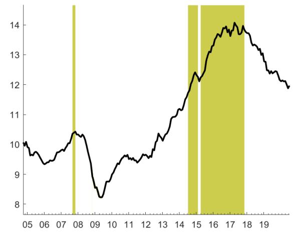


(l) *Price to income ratio for Hammersmith and Fulham*

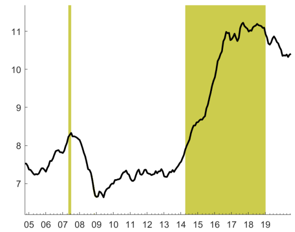
Figure 10: *Price to income ratio figures (i)*



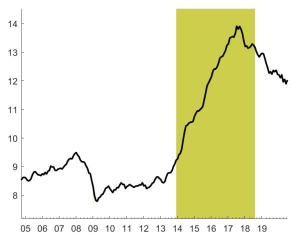
(a) *Price to income ratio for Haringey*



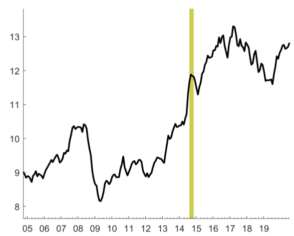
(b) *Price to income ratio for Harrow*



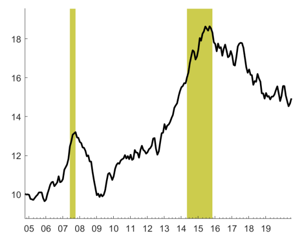
(c) *Price to income ratio for Havering*



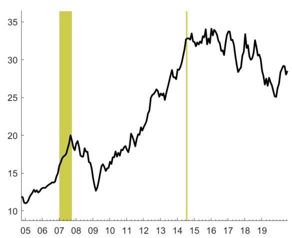
(d) *Price to income ratio for Hillingdon*



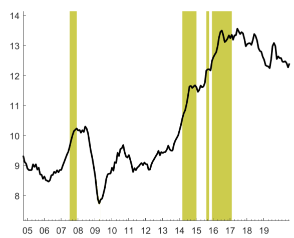
(e) *Price to income ratio for Hounslow*



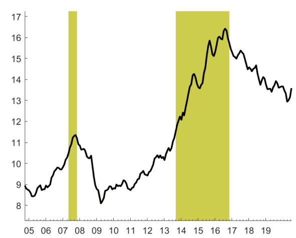
(f) *Price to income ratio for Islington*



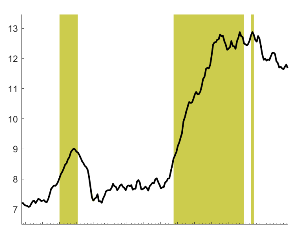
(g) *Price to income ratio for Kensington and Chelsea*



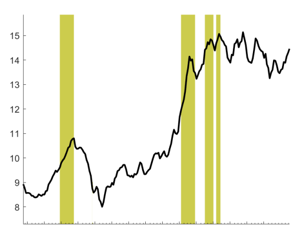
(h) *Price to income ratio for Kingston upon Thames*



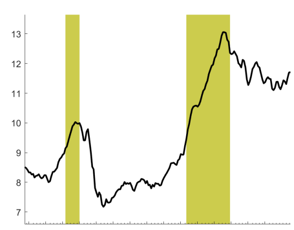
(i) *Price to income ratio for Lambeth*



(j) *Price to income ratio for Lewisham*

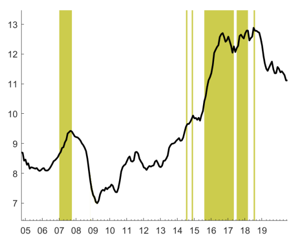


(k) *Price to income ratio for Merton*

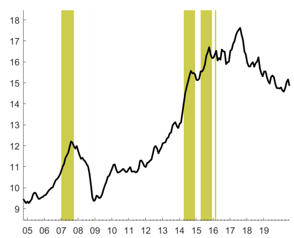


(l) *Price to Income ratio for Newham*

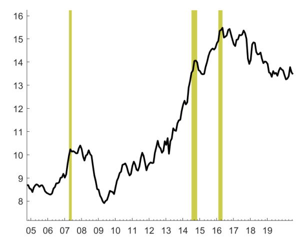
Figure 11: *Price to income ratio figures (ii)*



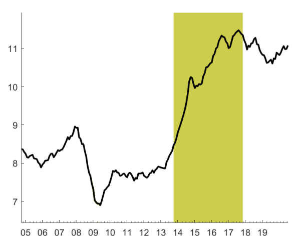
(a) *Price to income ratio for Redbridge*



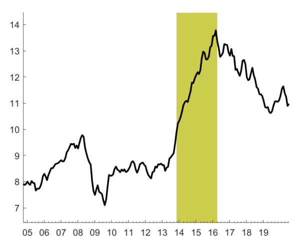
(b) *Price to income ratio for Richmond upon Thames*



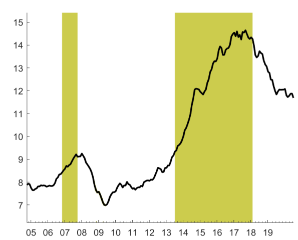
(c) *Price to income ratio for Southwark*



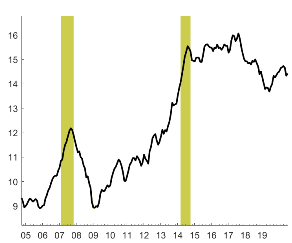
(d) *Price to income ratio for Sutton*



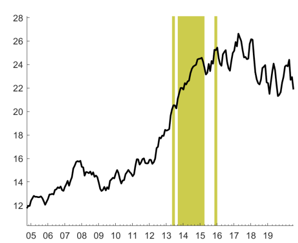
(e) *Price to income ratio for Tower Hamlets*



(f) *Price to income ratio for Waltham Forest*



(g) *Price to income ratio for Wandsworth*



(h) *Price to income ratio for Westminster*

Figure 12: *Price to income ratio figures (iii)*

Price to Income Ratio		
Borough	GSADF Statistic	Critical Value
Barking and Dagenham	**5.006044274**	1.396494327
Barnet	**4.4401648**	1.567470278
Bexley	**7.481367942**	1.555441685
Brent	**3.411305246**	1.522623449
Bromley	**6.396515851**	1.336529489
Camden	1.519096847	1.558319844
Croydon	**6.001139238**	1.587555859
Ealing	**4.41820265**	1.654518558
Enfield	**5.163598013**	1.537645403
Greenwich	**4.346006566**	1.457898813
Hackney	**3.900843004**	1.553891508
Hammersmith and Fulham	**2.540046119**	1.655567893
Haringey	**3.79179883**	1.45348224
Harrow	**3.184747451**	1.510935525
Havering	**6.554261625**	1.504893517
Hillingdon	**5.873862707**	1.639403163
Hounslow	**2.825262397**	1.882925875
Islington	**3.180793972**	1.518493338
Kensington and Chelsea	**4.154761828**	1.965076704
Kingston upon Thames	**4.557100844**	1.522623765
Lambeth	**4.052082706**	1.359102222
Lewisham	**4.502216034**	1.38467318
Merton	**5.090221777**	1.595453039
Newham	**4.867983458**	1.677145773
Redbridge	**4.93739562**	1.531524598
Richmond upon Thames	**5.184014165**	1.713125
Southwark	**3.083257277**	1.548628499
Sutton	**7.986116525**	1.610673724
Tower Hamlets	**3.360371124**	1.642958173
Waltham Forest	**7.112249645**	1.525971033
Wandsworth	**3.42312099**	1.484607603
Westminster	**2.878557125**	1.7005243
London	**3.994128569**	1.472540501

Table 1: *Price to income ratio GSADF statistics*

Again, the same point regarding the level of the ratio and overpricing with the price to income ratio still stands. The second bubble period identified ends in 2018. However, the level of the price to rent ratio is still much higher than the level identified in 2010. If housing is overpriced due to a bubble in 2010, why does it cease to be overpriced when the level is still higher in 2019?

Figures 14 and 15 show the results at the borough level. Generally speaking the results are far less uniform than with the price to income ratio. There are only seven regions where a bubble is identified in the financial crisis period: Brent, Camden & City of London, Lambeth, Lewisham & Southwark, Redbridge & Waltham Forest, Tower Hamlets, and Westminster. Compared to the price to income ratio, the boroughs of Camden, Westminster, and Tower Hamlets were all identified as having a bubble when testing the price to rent ratios but not the price to income ratios, during this period. Similarly, though many of the boroughs show a general upward trend in the price to rent ratio, there are only eight regions where a bubble is identified in the mid-2010s period with the price to rent ratio: Barking & Dagenham & Havering, Barnet, Bexley & Greenwich, Croydon, Enfield, Hackney & Newham, Merton, Kingston upon Thames & Sutton, and Redbridge & Waltham Forest. Again, the previous point about the level of the ratio applies again, even more so here. Brent, Lambeth, Tower Hamlets and Westminster are all identified as bubbles during the financial crisis period. However, for each of these regions, the price to rent ratio continues on a general upward trend, and ends up much higher that during the level which a bubble was identified as. GSADF test results are presented in Table 2. These largely mirror the BSADF results - however note that with the GSADF tests, Ealing tests positive for having explosive behaviour (the null of no explosive periods is rejected). However, this was excluded from

Figure 14 due to the only period identified being a collapsing bubble (collapsing bubbles being identified by the GSADF/BSADF procedure is discussed in Appendix vii).

Now if this were the final analysis done, it would be worth comparing and contrasting the results in more detail. For example, pinpointing the regions where exactly the bubbles took place/were centered in. It is possible that the bubble would be isolated to more affluent areas, for instance, which would be possible to identify given the regional level analysis. However, as I have explained, these two ratios ignore many factors which could account for changing house prices, primarily interest rates. Changes in interest rates are accounted for in the user cost to rent ratio, which are the next affordability ratios I test.

Price to Rent Ratio

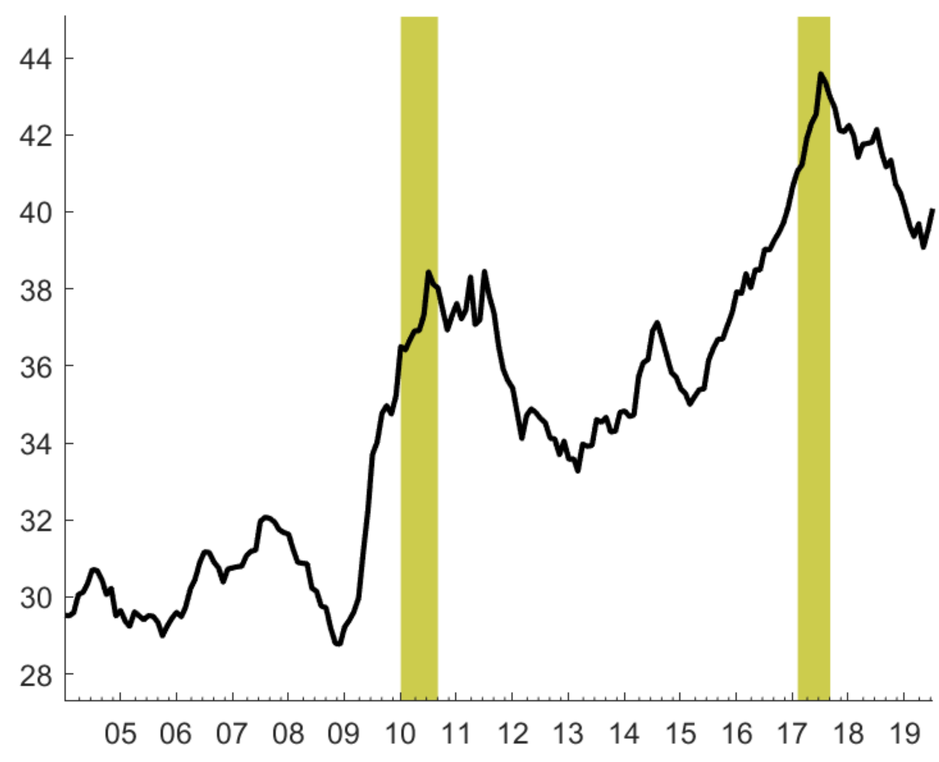
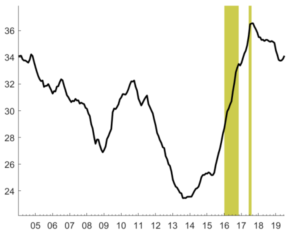


Figure 13: *Price to rent ratio for London*

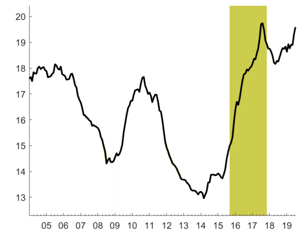
User Cost to Rent Ratio

Finally, I tested two variations of the user cost to rent ratio, with variation on how the expected price growth was calculated. Figure 16 shows the user cost to rent ratio for London, with the expected price growth calculated as a simple average of the last five years growth. Figure 19 shows the user cost to rent ratio for London, with the expected price growth calculated as a an optimised forecast using ARMA(p,q) models (The RMSE for different models, estimated over the period February 1995 to June 2001, are shown in Table 3). The series are generally quite different, showing that the expected price growth greatly affects the calculation of the user cost. However there are two main similarities between the two: i) for both series, at least one bubble is identified in the GFC period and the period immediately before it. ii) In the post GFC period, i.e. the 2010s, the user cost to rent ratio declined, and declined sharply. For both series, there is no bubble in the mid-2010s, despite this being a period characterised by very large house price increases, and a bubble being identified if using the price to rent and price to income ratios. This highlights the main issue that I am identifying in this paper. If one focuses only on the price to rent or price to income ratios, this ignores interest rates. And changing mortgage rates, caused by low interest rates, can explain changing house prices without there necessarily being a bubble. Now having made this point, it is worth noting that there still are bubbles identified using the user cost to rent ratio. Furthermore this is a result that cannot be dismissed with appeal to other factors, in the same way that the price to rent/income ratios can be dismissed by appeal to changing interest rates. The identification of a bubble in the GFC period using the user cost to rent ratio appears very robust, given that one cannot appeal to other factors. Most major factors enter the user cost to rent directly, and all factors that do not enter the ratio explicitly, e.g.



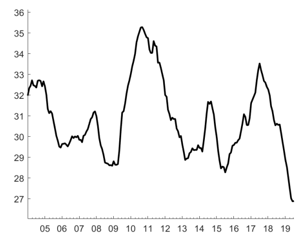
(a) *Price to rent ratio for Barking and Dagenham and Havering*

(b) *Price to rent ratio for Barnet*

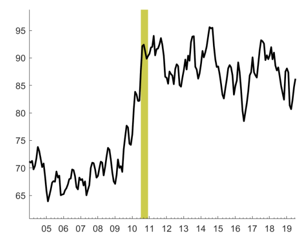


(c) *Price to rent ratio for Bexley and Greenwich*

(d) *Price to rent ratio for Brent*



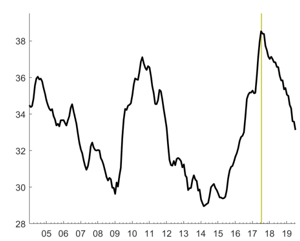
(e) *Price to rent ratio for Bromley*



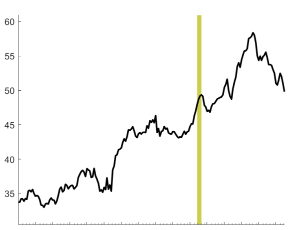
(f) *Price to rent ratio for Camden and City of London*

(g) *Price to rent ratio for Croydon*

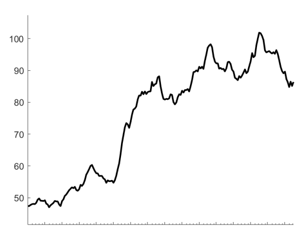
(h) *Price to rent ratio for Ealing*



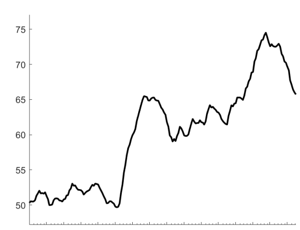
(i) *Price to rent ratio for Enfield*



(j) *Price to rent ratio for Hackney and Newham*

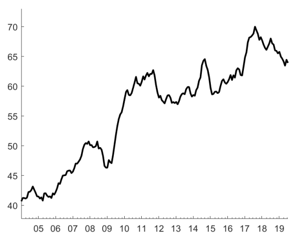


(k) *Price to rent ratio for Haringey and Islington*



(l) *Price to rent ratio for Harrow and Hillingdon*

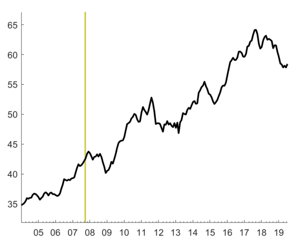
Figure 14: *Price to rent ratio figures (i)*



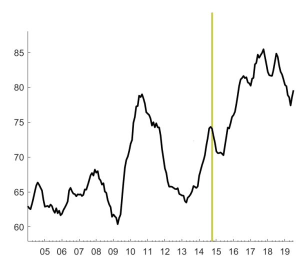
(a) *Price to rent ratio for Hounslow and Richmond upon Thames*

(b) *Price to rent ratio for Kensington and Chelsea and Hammersmith and Fulham*

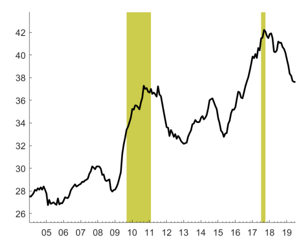
(c) *Price to rent ratio for Lambeth*



(d) *Price to rent ratio for Lewisham and Southwark*



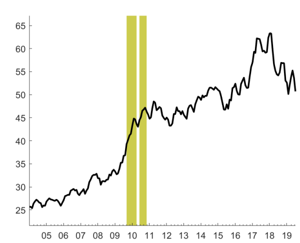
(e) *Price to rent ratio for Merton, Kingston upon Thames and Sutton*



(f) *Price to rent ratio for Redbridge and Waltham Forest*

(g) *Price to rent ratio for Tower Hamlets*

(h) *Price to rent ratio for Wandsworth*



(i) *Price to rent ratio for Westminster*

Figure 15: *Price to rent ratio figures (ii)*

Price to Rent Ratio		
Region	GSADF Statistic	Critical Value
Barking and Dagenham and Havering	**6.672681119**	4.393491142
Barnet	**2.718983045**	1.692492818
Bexley and Greenwich	**4.884533809**	1.549471546
Brent	**2.260150756**	1.65903549
Bromley	2.941903143	3.571104094
Camden and City of London	**3.792589636**	2.666971479
Croydon	**4.676636625**	1.339134065
Ealing	**1.43412027**	1.398786184
Enfield	**3.717678794**	3.693615361
Hackney and Newham	**3.108248059**	1.808143281
Haringey and Islington	2.625513291	3.186112829
Harrow and Hillingdon	3.006864502	3.312221005
Hounslow and Richmond upon Thames	2.594480805	2.602298599
Kensington and Chelsea and Hammersmith and Fulham	2.010560075	3.127073334
Lambeth	**1.916868256**	1.706567585
Lewisham and Southwark	**2.489570656**	2.433306183
Merton Kingston upon Thames and Sutton	**4.24286579**	3.818677608
Redbridge and Waltham Forest	**3.003702705**	1.523500447
Tower Hamlets	**2.846598513**	2.290131942
Wandsworth	2.849802265	2.996341225
Westminster	**4.089453853**	2.041425761
London	**3.043334686**	1.639104321

Table 2: *Price to rent ratio GSADF statistics*

income, should affect rental prices as well as house prices. If house prices are high due to a genuine lack of housing, this would also affect rental prices (Mulheirn, 2019). This result is robust to different model specifications - i.e. there is the same result using the five year average, and the optimised ARMA model forecast. The results are also unaffected by different methods to calculate the risk premium - as mentioned previously, results using alternate specifications with different risk premiums are presented in Appendix vii, and the results (in terms of where and when bubbles are identified) are almost identical [22]. This means that the bubble we *do* identify using the user cost to rent ratio is essentially indisputable.

Note however that there are some key differences between the two series. Using the five year average, the user cost to rent ratio identifies a bubble towards the end of the sample i.e. in 2019. It is very easy to provide an explanation of this given the data - house prices had been increasing for years, however in a way that was justified by the fundamentals (i.e. low interest rates). But what matters in the formation of bubbles is the expectation of future price increases. Hence because it was not properly understood that lower interest rates could cause house prices to increase, there was a perception of house prices constantly going up in a bubble like manner, which in a self-fulfilling manner precipitated the formation of a bubble towards the end of the sample. However, this 2019 bubble is not identified when calculating the user cost to rent ratio using the optimised forecasts for expected price growth. This shows how key price growth expectations are in calculating the user cost. Furthermore, there are differences in the general level of the two series. As I have

[22] Note however that the *level* of the ratio can be shifted by different risk premiums, which affects the timing of under and overpricing. This is discussed more in Appendixvii.

mentioned previously, we would expect the user cost to rent ratio to be around one, which means that the user cost and rent are equal to each other. If this were not so, there would be disequilibrium between renting and owner-occupying housing. We would not expect it to be exactly one all of the time, because as I mentioned in the literature review, the housing market is characterised by frictions that mean it is unlikely to be efficient in the same way the stock market is (i.e. returns being a martingale difference sequence). It might differ significantly from one for extended periods of time, before returning to one. With the user cost calculated with the five year average, this is generally what I find. There is a high degree of persistence in the series, however when the ratio is much higher or lower than one, it will stay this way for no more than several years before reversing direction. At the start of the series it is *negative*, which means that housing is essentially a free lunch at this point - but it quickly rises to around one for a few years, until 2008 where it starts rising rapidly. It peaks at four in 2009, it maintains this level for around three years, before collapsing around 2013 (even though nominal house prices were still increasing at this point, the user cost of housing was actually *decreasing* because of low interest rates). After this, the user cost to rent ratio becomes very low (going negative briefly in 2016 and again in 2017) before starting to rise again, at which point in 2019 it is above one again. One thing to reiterate is that using this ratio gives us a clear idea of when exactly the housing market is over or underpriced. When the ratio is above one, housing is overpriced. When the ratio is below one, it is underpriced. For instance, it could be argued that the bubble identified at the start of the sample is not a bubble at all, but is actually just a rapid correction from the housing market being under-priced (though it does continue to be overpriced, and a second period of explosiveness is identified). One could describe this oscillation observed as the housing market going through distinct periods of under- and overpricing. This is in line with the behavioural finance literature, which suggests that the under and overpricing periods correspond to under and overconfidence respectively (Barberis et al., 1998; Daniel et al., 1998; Shiller, 2000). The expectations of the market become too high during booms, which leads to overpricing, but when the market crashes, the expectations of the market become underconfident and excessively negative, which leads to underpricing. When calculating the expected price growth using optimised forecasts, the level of the user cost ratio is totally different. Generally speaking, the level is consistently much higher than one - even when it declines, it does not decline to a level below one. This indicates that when the user cost is calculated this way, housing is overpriced for the entire sample. This seems unrealistic, and not what we would expect. Perhaps this is evidence in favour of using the five year average for the expected price growth - given the unrealistic results obtained from using the forecasts.

Figures 17 and 18 show the results using the five year average as expected price growth at the regional level. Generally the results are very similar to London overall, for all the boroughs. All of the boroughs apart from Westminster have a bubble identified in the pre-2010 period, though the exact timing varies. None of the boroughs have a bubble identified during the mid-2010 period. Eleven of the regions do have a bubble identified near the end of the sample, i.e. around 2019. These are Barking & Dagenham & Havering, Barnet, Brent, Croydon, Enfield, Haringey & Islington, Harrow & Hillingdon, Lambeth, Lewisham & Southwark, Merton & Kingston upon Thames & Sutton, and Redbridge & Waltham Forest. Though even in the regions where explosiveness was not identified, the user cost to rent ratio does increase in 2019 for all of these boroughs. The point about the level of the user cost ratio oscillating around one can be reiterated at the regional level. Every single borough goes through periods of over and underpricing, though the exact magnitude of the over and underpricing varies for all of the regions. For instance, in Kensington & Chelsea & Hammersmith & Fulham, the user cost to rent ratio reaches *fifteen* in 2012, before collapsing to slightly negative in 2016. By contrast, the user cost to rent ratio in Croydon is bounded by approximately two and negative one. So while there does not seem to be a particular trend on if the bubble took place in the more expensive vs less expensive regions, there definitely does seem to be some differences in the magnitude of the overpricing, given that Kensington & Chelsea & Hammersmith & Fulham are two of the most expensive boroughs in London, and also had the highest levels of overpricing. GSADF test results are reported in Table 4, and largely mirror the results from the BSADF procedure (Westminster tests positive with the GSADF test as having explosive behaviour, however the only period identified for Westminster by the BSADF procedure was a collapsing bubble).

Figures 20 and 21 show the results at the regional level, using the optimised ARMA model forecasts for expected price growth. There are more regional variations using this version of the user cost. There are only six regions with a bubble identified in the pre-2010 period - Brent, Haringey & Islington, Kensington & Chelsea & Hammersmith & Fulham, Redbridge & Waltham Forest, Tower Hamlets, and Wandsworth. However, there *is* a bubble identified in the mid-2010s in one region - Haringey & Islington (which is one of the more expensive regions of London). Like with the user cost calculated with the expected price growth as an average, all of the regions show a general downward trend in the user cost relative to rent in the period post-GFC. Which means that however the expected price growth is calculated, the user cost to rent ratio declined. House prices may have been increasing in nominal terms, however this was justified by the fundamentals. In general the point about the levels of the ratio being unrealistically high when calculated with the forecast instead of the average for the expected price growth applies here again. Most of the regional user cost to rent ratios with the forecast show consistent massive overpricing, with the level much greater than one for the entire sample. There are some exceptions to this however - some of the regions start with a large level of overpricing, but end the sample with the ratio being approximately one. This happens in Bexley & Greenwich, Bromley, Croydon, Enfield, Hounslow & Richmond upon Thames, Kensington & Chelsea & Hammersmith and Fulham, Lewisham & Southwark, Westminster. Furthermore the ratio in Lambeth starts around one, but becomes negative by the end of the sample. GSADF test results are reported in Table 5. With the GSADF tests, Barking & Dagenham & Havering, Bexley & Greenwich, Bromley, Croydon, Enfield, Lambeth, Lewisham & Southwark test positive for having explosive behaviour, however the only periods identified by the BSADF procedure were for a collapsing bubble.

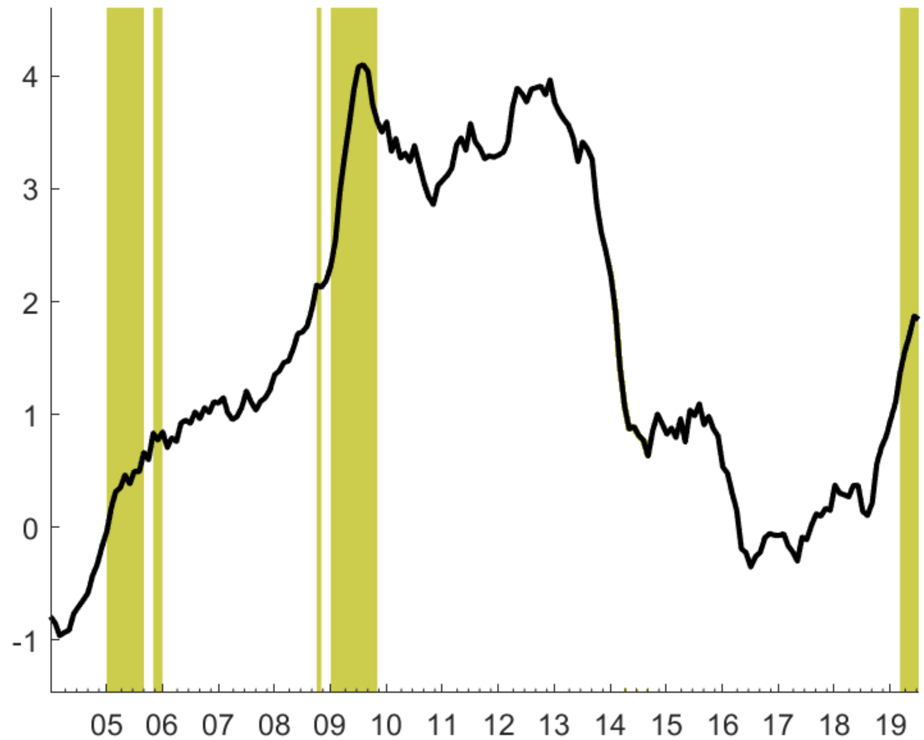
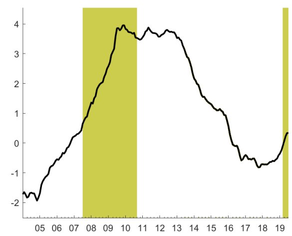
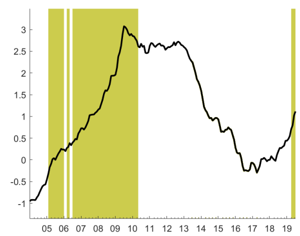


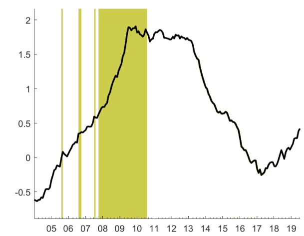
Figure 16: *User cost to rent ratio for London (with five year average expected price growth)*



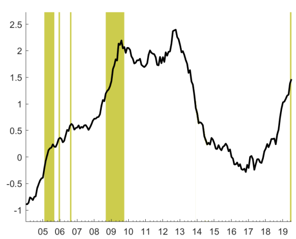
(a) *User cost to rent ratio for Barking and Dagenham and Havering*



(b) *User cost to rent ratio for Barnet*

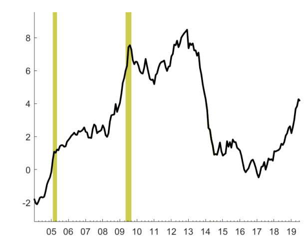


(c) *User cost to rent ratio for Bexley and Greenwich*



(d) *User cost to rent ratio for Brent*

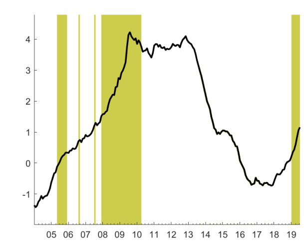
(e) *User cost to rent ratio for Bromley*



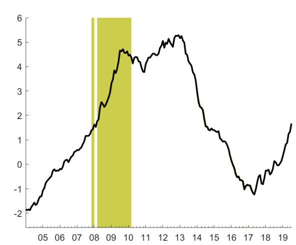
(f) *User cost to rent ratio for Camden and City of London*

(g) *User cost to rent ratio for Croydon*

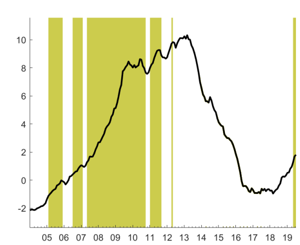
(h) *User cost to rent ratio for Ealing*



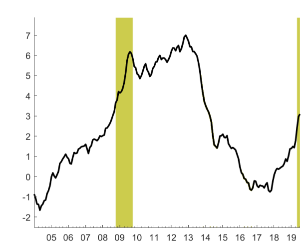
(i) *User cost to rent ratio for Enfield*



(j) *User cost to rent ratio for Hackney and Newham*

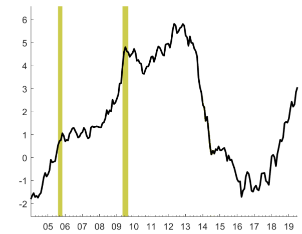


(k) *User cost to rent ratio for Haringey and Islington*

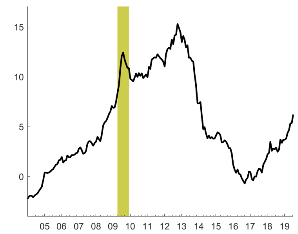


(l) *User cost to rent ratio for Harrow and Hillingdon*

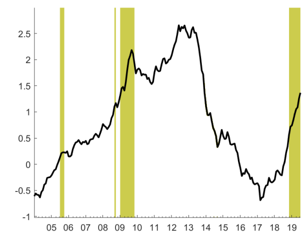
Figure 17: *User cost to rent ratio figures (with five year average expected price growth) (i)*



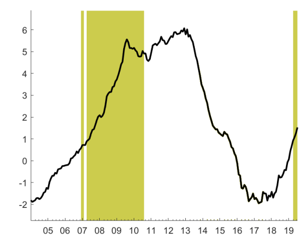
(a) *User cost to rent ratio for Hounslow and Richmond upon Thames*



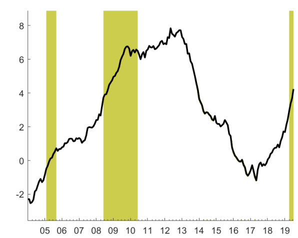
(b) *User cost to rent ratio for Kensington and Chelsea and Hammersmith and Fulham*



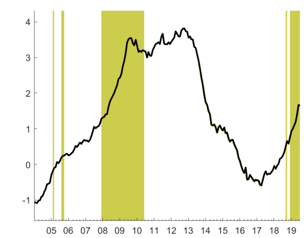
(c) *User cost to rent ratio for Lambeth*



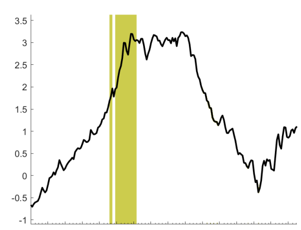
(d) *User cost to rent ratio for Lewisham and Southwark*



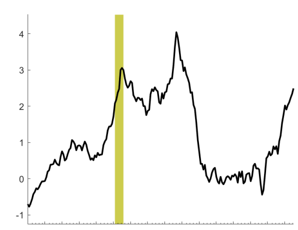
(e) *User cost to rent ratio for Merton, Kingston upon Thames and Sutton*



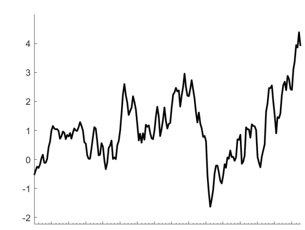
(f) *User cost to rent ratio for Redbridge and Waltham Forest*



(g) *User cost to rent ratio for Tower Hamlets*



(h) *User cost to rent ratio for Wandsworth*



(i) *User cost to rent ratio for Westminster*

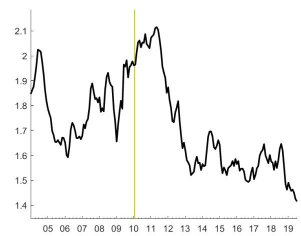
Figure 18: *User cost to rent ratio figures (with five year average expected price growth) (ii)*

Figure 19: *User cost to rent ratio for London (with optimised forecast expected price growth)*

(a) *User cost to rent ratio for Barking and Dagenham and Havering*

(b) *User cost to rent ratio for Barnet*

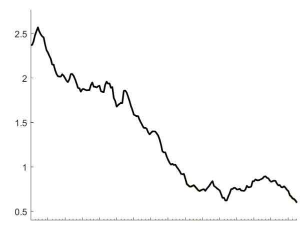
(c) *User cost to rent ratio for Bexley and Greenwich*



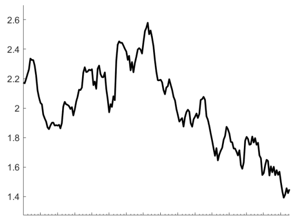
(d) *User cost to rent ratio for Brent*

(e) *User cost to rent ratio for Bromley*

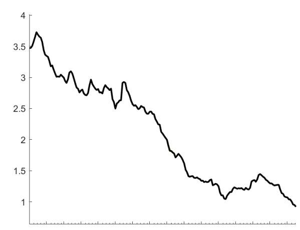
(f) *User cost to rent ratio for Camden and City of London*



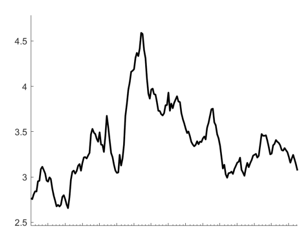
(g) *User cost to rent ratio for Croydon*



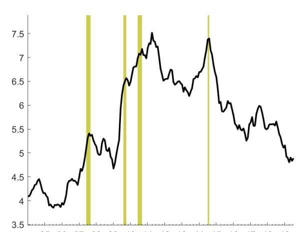
(h) *User cost to rent ratio for Ealing*



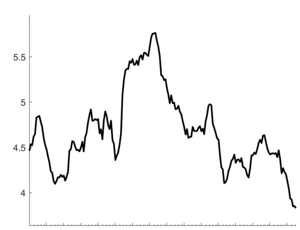
(i) *User cost to rent ratio for Enfield*



(j) *User cost to rent ratio for Hackney and Newham*



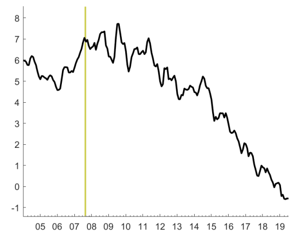
(k) *User cost to rent ratio for Haringey and Islington*



(l) *User cost to rent ratio for Harrow and Hillingdon*

Figure 20: *User cost to rent ratio figures (with optimised forecast expected price growth) (i)*

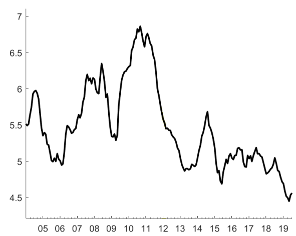
(a) *User cost to rent ratio for Hounslow and Richmond upon Thames*



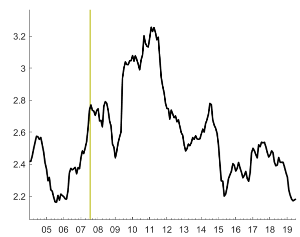
(b) *User cost to rent ratio for Kensington and Chelsea and Hammersmith and Fulham*

(c) *User cost to rent ratio for Lambeth*

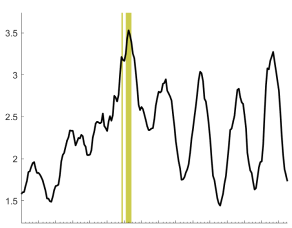
(d) *User cost to rent ratio for Lewisham and Southwark*



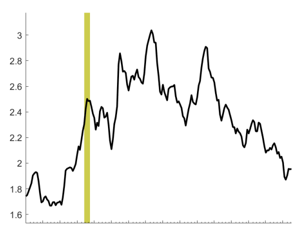
(e) *User cost to rent ratio for Merton, Kingston upon Thames and Sutton*



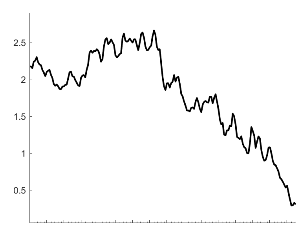
(f) *User cost to rent ratio for Redbridge and Waltham Forest*



(g) *User cost to rent ratio for Tower Hamlets*



(h) *User cost to rent ratio for Wandsworth*



(i) *User cost to rent ratio for Westminster*

Figure 21: *User cost to rent ratio figures (with optimised forecast expected price growth) (ii)*

	Root Mean Square Error (RMSE)					
Region/Forecast model	AR(1)	AR(2)	MA(1)	MA(2)	ARMA(1,1)	ARMA(2,2)
Barking and Dagenham and Havering	**0.036804**	0.036935	0.135572	0.053303	0.051231	0.051726
Barnet	**0.096747**	0.105306	0.134566	0.122225	0.116626	0.149097
Bexley and Greenwich	**0.05839**	0.065105	0.134711	0.071099	0.096997	0.098801
Brent	0.070184	**0.066658**	0.135654	0.123861	0.080748	0.104762
Bromley	0.102353	0.104176	0.141388	0.129722	**0.101935**	0.120542
Camden and City of London	**0.077021**	0.077989	0.104766	0.096194	0.122357	0.104708
Croydon	0.143594	0.144915	0.144903	**0.132134**	0.177852	0.177095
Ealing	0.104972	**0.081589**	0.108364	0.099172	0.123011	0.130696
Enfield	0.081554	0.084178	0.09486	**0.075608**	0.117771	0.111906
Hackney and Newham	0.072463	**0.067797**	0.182859	0.167455	0.119663	0.107161
Haringey and Islington	**0.054533**	0.059047	0.133638	0.122321	0.077898	0.074288
Harrow and Hillingdon	0.083964	**0.078721**	0.114443	0.103649	0.117498	0.115185
Hounslow and Richmond upon Thames	0.113761	0.094181	0.131453	0.118792	0.073482	**0.04125**
Kensington and Chelsea and Hammersmith and Fulham	0.088829	0.088378	0.074752	0.069178	**0.051924**	0.053169
Lambeth	0.058575	0.056864	0.114549	0.107594	**0.044402**	0.048616
Lewisham and Southwark	0.05815	0.059007	0.182159	0.165737	**0.045758**	0.046056
Merton, Kingston upon Thames and Sutton	**0.098152**	0.112541	0.121853	0.112148	0.109869	0.160527
Redbridge and Waltham Forest	0.134197	0.166582	0.13384	**0.122743**	0.159925	0.207443
Tower Hamlets	0.058928	**0.055342**	0.071217	0.066622	0.057326	0.063435
Wandsworth	0.138855	0.141978	0.140887	**0.129793**	0.136118	0.217941
Westminster	**0.058023**	0.064149	0.089072	0.083884	0.066627	0.104595
London	**0.077914**	0.079135	0.111514	0.102218	0.12682	0.12521

Table 3: *Root mean square error (RMSE) for different models*

User Cost to Rent Ratio (five year average)		
Region	GSADF Statistic	Critical Value
Barking and Dagenham and Havering	**5.56630697**	1.658417574
Barnet	**6.565163944**	1.243649511
Bexley and Greenwich	**4.690264461**	1.395244092
Brent	**3.589751381**	1.511854967
Bromley	**2.348615178**	1.636960513
Camden and City of London	**3.412289976**	1.608646619
Croydon	**2.560010034**	1.504027082
Ealing	**2.741609567**	1.494274553
Enfield	**5.465202604**	1.532481551
Hackney and Newham	**4.130894366**	1.621902591
Haringey and Islington	**8.04364677**	1.727467849
Harrow and Hillingdon	**4.07194899**	1.653552934
Hounslow and Richmond upon Thames	**3.4370424**	1.586403341
Kensington and Chelsea and Hammersmith and Fulham	**4.854496244**	1.671396162
Lambeth	**3.549002883**	1.552040564
Lewisham and Southwark	**5.622590591**	1.504786204
Merton Kingston upon Thames and Sutton	**3.946177495**	1.540786126
Redbridge and Waltham Forest	**5.916927054**	1.249184694
Tower Hamlets	**3.884989597**	1.608445721
Wandsworth	**3.464567822**	1.53228459
Westminster	**1.900990073**	1.650467272
London	**6.204246931**	1.464099644

Table 4: *User cost to rent ratio (five year average) GSADF Statistics*

User Cost to Rent Ratio (optimised forecast)		
Region	GSADF Statistic	Critical Value
Barking and Dagenham and Havering	**3.075007251**	1.693887
Barnet	1.506393839	1.596727
Bexley and Greenwich	**2.275037313**	1.468211
Brent	**1.691343248**	1.499913
Bromley	**1.905569502**	1.524897
Camden and City of London	0.594257372	1.408537
Croydon	**2.295007546**	1.581116
Ealing	1.170181353	1.548727
Enfield	**2.624111831**	1.51431
Hackney and Newham	1.412898423	1.416636
Haringey and Islington	**2.54833133**	1.553331
Harrow and Hillingdon	1.470745374	1.485801
Hounslow and Richmond upon Thames	1.703324218	1.80891
Kensington and Chelsea and Hammersmith and Fulham	**1.628013126**	1.450327
Lambeth	**1.835235074**	1.394005
Lewisham and Southwark	**1.693211579**	1.65254
Merton Kingston upon Thames and Sutton	**2.017806818**	1.690602
Redbridge and Waltham Forest	**2.161311419**	1.574978
Tower Hamlets	**2.358839577**	1.397395
Wandsworth	**2.990607598**	1.547592
Westminster	1.187113002	1.42053
London	**2.372343039**	1.422788

Table 5: *User cost to rent ratio (optimised forecast) GSADF Statistics*

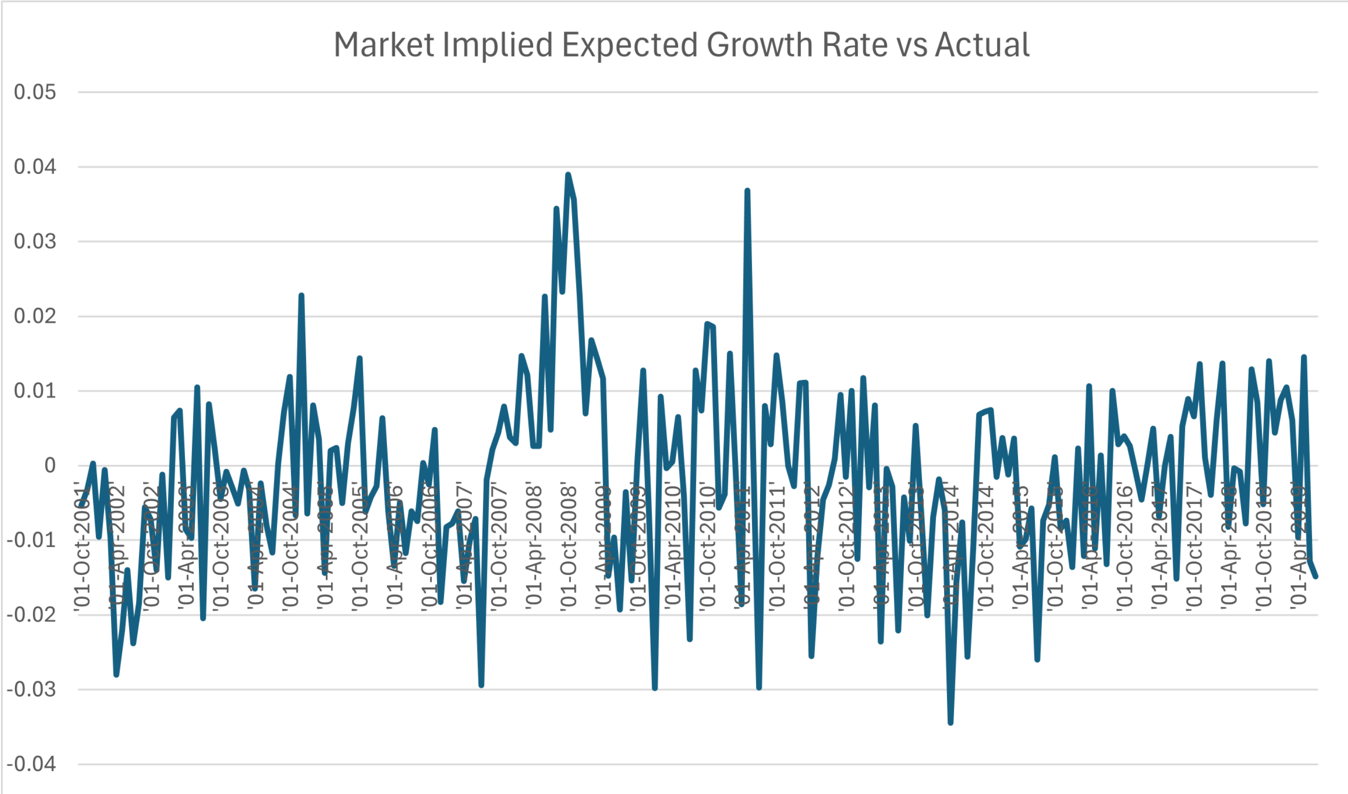


Figure 22: *Market implied expected growth rate minus actual growth rate*

vi. Conclusion and Areas for Further Research

In the Introduction, I noted how house prices increased at a high pace in the 2000s and 2010s. I have shown that if one tests the house price to income and house price to rent ratios with a PSY recursive unit root test procedure, one would identify two primary bubble periods - around the GFC, and in the mid-2010s. However, if one tests the user cost to rent ratio, which incorporates the effect of changing interest rates and other factors into housing affordability, there is no bubble during the mid-2010s. This means that despite sharp house price rises in the post-GFC period, the price rises were explainable by changes in interest rates. While nominal prices were increasing, the actual cost of home ownership was decreasing. While the analysis focused on a situation with mortgage rates, where the mortgage interest payments and other costs of home ownership should be equal to the cost of renting, as I made clear in the Introduction and Literature Review, the same applies to a cash buyer as well, given that the same dynamics that affect mortgage rates will also affect the discount rate for a cash-buyer[23].

Since 2019, there have been numerous shocks to the housing market. The primary shock that is currently of huge importance are changes to interest rates. Since 2022, interest rates, and mortgage rates, have started to increase. This is coming from a time where mortgage rates were at record lows. The theory, and results of this paper, would suggest that a house price collapse is imminent, as housing becomes unaffordable. After all, if house prices will go up because interest rates have gone down, surely it is also the case that housing prices will go down if interest rates go up. This is not necessarily the case however. The primary reason that a homeowner would sell their house with a massive capital loss is where there are arrears with mortgages, and the house is repossessed. During the 1990s housing crash, there was a large wave of mortgage repossessions, and house prices plummeted[24]. However in recent years there have been substantive efforts by governments to limit the

[23]Now it is worth noting that while it is true that for a typical homeowner buying with a mortgage or cash, housing was generally not overpriced in the post-GFC period, this does not mean that there are not housing problems in the London housing market, and the UK housing market more generally. There has been a lack of new social housing built in the UK, and the rental market is relatively unregulated (for instance, compared to countries like Germany) (Mulheirn, 2019).

[24]In 1991, there were 75,500 mortgage repossessions (see Table 1300 on https://www.gov.uk/government/statistical-data-sets/live-tables-on-repossession-activity). This compares to a mere 4,620 in 2023 (See Table 1 on Mortgage and landlord possession statistical tables, accessed via https://www.gov.uk/government/statistics/mortgage-and-landlord-possession-statistics-january-to-march-2024)

amount of foreclosures, given the unaffordability of current mortgage rates. In 2023 the UK government introduced the Mortgage Charter, which set out a number of measures with the aim of avoiding a large wave of repossessions[25]. It would be politically unattractive for a government to allow a housing crash to occur, so it is likely that future UK governments, and governments across the world, will introduce further measures in the future to mitigate foreclosures if the need arises. Obviously this could still occur, but in general if there is no large level of foreclosures (or other forced sales of properties), then homeowners who do not wish to sell at a massive loss do not need to do so, and thus there would be no crash. Housing is a durable asset, which can simply be left empty. What would happen instead is that houses would simply not change hands.

And this is what has been happening since 2022. At the time of writing, the most recent estimate for monthly property transactions in the UK (79,590, April 2024) is down 17% since two years ago (99,510 in April 2022)[26].

As the housing market is out of equilibrium, and homeowners are not compelled to sell, house prices are too high given the economic fundamentals, but because homeowners do not wish to sell, they can simply not sell, forcing transactions to plummet. There are several ways that the housing market could return to equilibrium. One possibility is that interest rates return to their pre-COVID levels before housing prices crash. I find this incredibly unlikely however, given that the general consensus seems to be that the central bank rate is going to decrease to approximately 3.75% by the end of 2025. Before 2022, *mortgage*

[25] https://www.gov.uk/government/publications/mortgage-charter/mortgage-charter

[26] According to 'UK monthly property transactions tables' accessed via https://www.gov.uk/government/statistics/monthly-property-transactions-completed-in-the-uk-with-value-40000-or-above, accessed 17/06/2024

rates (which are generally several percentage points above the central bank rate) were lower than this. Another possibility for the housing market returning to equilibrium is that housing prices could crash - though despite unaffordable mortgage rates, this is only likely to happen if there are a large number of forced sales. A more likely scenario is that house prices slowly adjust to a new level - especially if this were in real terms, due to inflation. Consider the following: A common finding from the behavioural finance literature is that people tend to "anchor" prices[27], and many people have a psychological aversion to loss[28], such that they are unwilling to sell at prices below that which they bought at. For example, suppose someone bought a house for £300,000, but economic conditions shift (e.g. interest rates go up) such that the market clearing price becomes £250,000. Many people would be unwilling to sell at less than £300,000 given that this entails a capital loss. However, suppose that inflation were 4%, for five years. Now, £300,000 would be equivalent to £250,000 five years prior, in real terms. Now there would be no nominal loss on the sale if it were sold for £300,000, even though in real terms there was. Hence such a person would likely be more willing sell, even though in real terms the loss is the same. In the 1990s housing crash, there were several years of steep losses in nominal terms. However, in *real terms*, housing prices took many years to reach a plateau. Figure 23 shows how UK house prices collapsed in the 1990s. In nominal terms, UK house prices took approximately three years to decrease from £60,701 in 1989Q3 to £53,213 in 1992Q4, a decline of 12.3%. However in real terms, house prices took an additional four years to reach a plateau of

[27]The seminal paper that introduced anchoring, and alluded to this happening in an economic context, is Tversky and Kahneman (1974). An early example in the literature which finds experimental evidence of anchoring in a real world context is Northcraft and Neale (1987).

[28]Some seminal works in this literature are Tversky and Kahneman (1979), Tversky and Kahneman (1992).

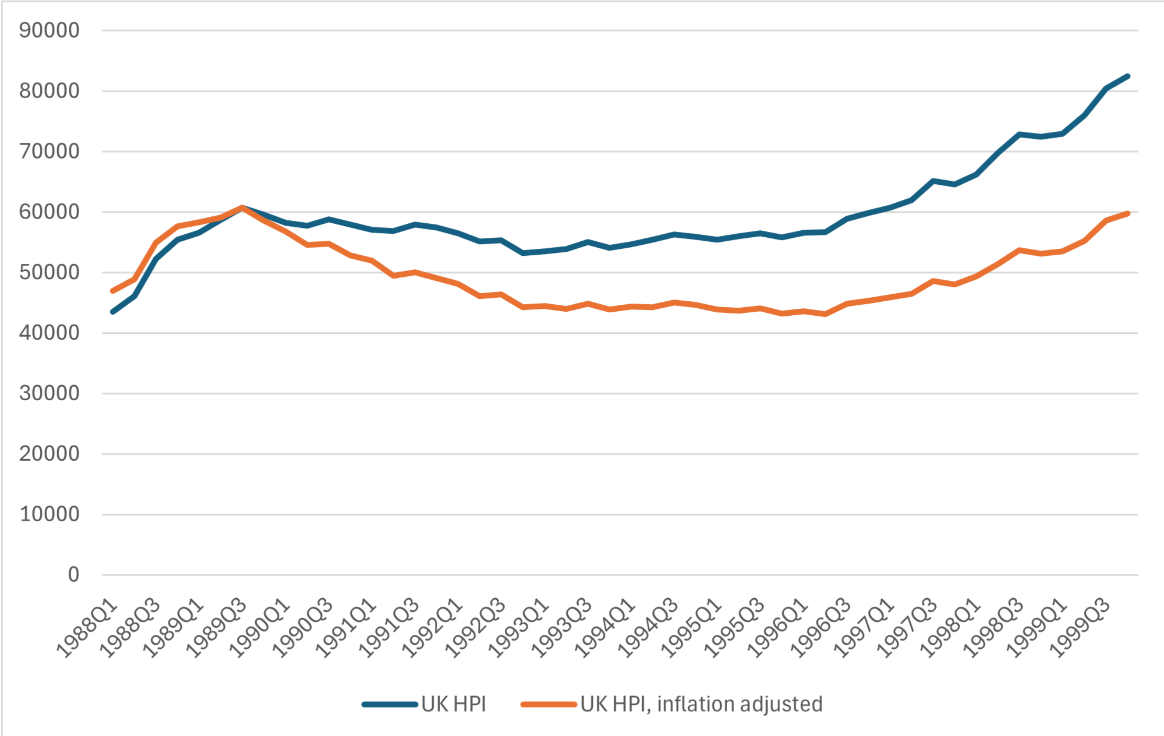


Figure 23: *The 1990s UK housing crash, in nominal terms and inflation adjusted*

£43,150 (in 1989Q3 pounds) in 1996Q2, a collapse in real prices of 28.9%. In real terms, it took approximately ten years for house prices to return to their 1989 level.

Now one of the key issues that has been highlighted in this paper is how significant the expected price growth is for calculating the user cost. While both methods for calculating the expected price growth resulted in very generally the same conclusions (i.e. identification of a bubble in the GFC period, and no bubble (and in fact declining user cost, relative to rent) in the post GFC period in spite of nominal price increases), the level of the ratios, and thus specifically whether housing is over or underpriced, is totally different depending on how the expected price growth is calculated. This has been noted previously in the literature: there is, however, a potential way around this, that has been only very briefly touched upon (Hill and Syed (2016) provides a good discussion of this topic). In theory, the user cost should be equal to rent (at least over the long term). So the following equation should hold:

$$R_t = P_t(r_{m,t} + \delta_t + \tau_{p,t} + \alpha_t - E(\dot{P}_t)) \tag{20}$$

Every single factor in this equation is either directly observable, or there is a good proxy, except for the risk premium and expected price growth. A simple solution then, is to suppose: Assume that this relationship holds, what would the expected price growth and risk premium have to be for the rent to be equal to the user cost? The formula can be simply rearranged, to get the unobserved terms on the LHS of the equation:

$$E(\dot{P}_t) - \alpha_t = r_{m,t} + \delta_t + \tau_{p,t} - \frac{R_t}{P_t} \tag{21}$$

What we have on the LHS is the *market implied* expected price growth, minus the risk premium. As I have noted, different methods for calculating the risk premium had a much smaller effect on the results than different methods for the expected price growth. It is generally standard in the literature to assume that it is constant. In this paper, the estimates for the risk premium using conventional methods (the CAPM and CCAPM) resulted in a risk premium that was extremely small, sometimes negative. In practice, one could assume that it is either zero or fixed. What one could then do is compare the market implied expected price growth to actual price growth. For convenience, I refer to this as the forecast error: how much higher the market implied growth rate is than the actual growth rate. This forecast error could then give an indication of where a bubble is. If the forecast error is significantly higher than zero this would show that there are biased expectations about future price growth, which is ultimately one of the main hallmarks of an asset price bubble.

I actually plot this in Figure 22, assuming that the risk premium is zero[29]. Now it does visually look like the forecast error was much higher than zero around the GFC. It is worth noting however, that the current econometric framework (i.e. recursive explosiveness tests) would not be appropriate here, as we would expect forecast errors to be a stationary process. I think this is a promising area for future research, however an alternative econometric framework would be required (perhaps recursive *stationarity* tests instead).

vii. Appendix

Collapsing Prices as Explosive

The BSADF procedure has a tendency to highlight collapsing prices as explosive, as well as explosive increases. To understand why this happens, first note that the ADF test is essentially an OLS regression, with the parameter of interest being the slope parameter, δ, in the following equation:

$$\Delta y_t = \mu + \delta y_{t-1} + \sum_{i=1}^{k} \alpha_i \Delta y_{t-i} + \varepsilon_t \tag{22}$$

Furthermore, note that in OLS, the estimate of the slope parameter is invariant to the level of the points. A shift in the level of the data points won't affect the value of the slope. This is illustrated in Figure 24.

[29] In practice one could calibrate the risk premium to be the mean value of the forecast error over the whole sample, if this is significantly different from zero.

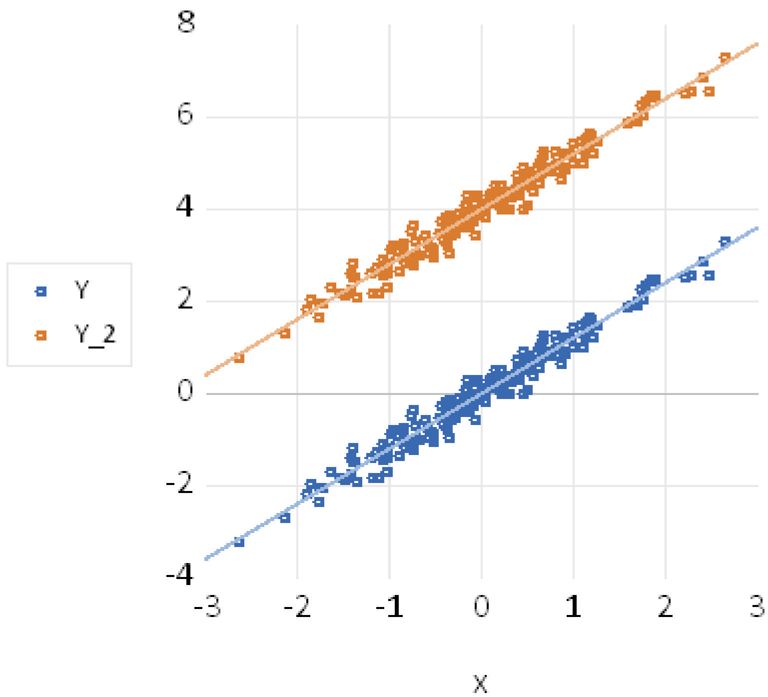


Figure 24: *A generic scatter plot: shifting up the points does not affect the value of the slope*

This is because the slope parameter is determined by the covariances and variances of the dependent and independent variables. If all the values of y_t are shifted up or down by some value a, the values of Δy_t are unchanged because these are determined by the differences in y_t. The value of the covariances of $Cov(y_t, ...)$ will be also be unchanged because, in general, for any variables X and Y and constants a and b, it is true that $Cov(a+X, b+Y) = Cov(X,Y)$. Hence $Cov(y_t + a, ...) = Cov(y_t, ...)$ for any constant a. So the value of the estimated slope parameter is invariant to the level of the data points, what matters is the relative position of the data points (assuming that there is an intercept term in the OLS equation).

What do explosive processes look like? In general, we associate explosive processes with an exponential increase. As the initial value is multiplied by some number greater than one, it gets bigger and bigger as time goes on. However, this only occurs if the initial value of the process is positive. If the initial value of the process is *negative*, then the value gets *smaller and smaller* as time goes on (though the magnitude does get larger in the same way). So both of the processes in Figure 25 are explosive.

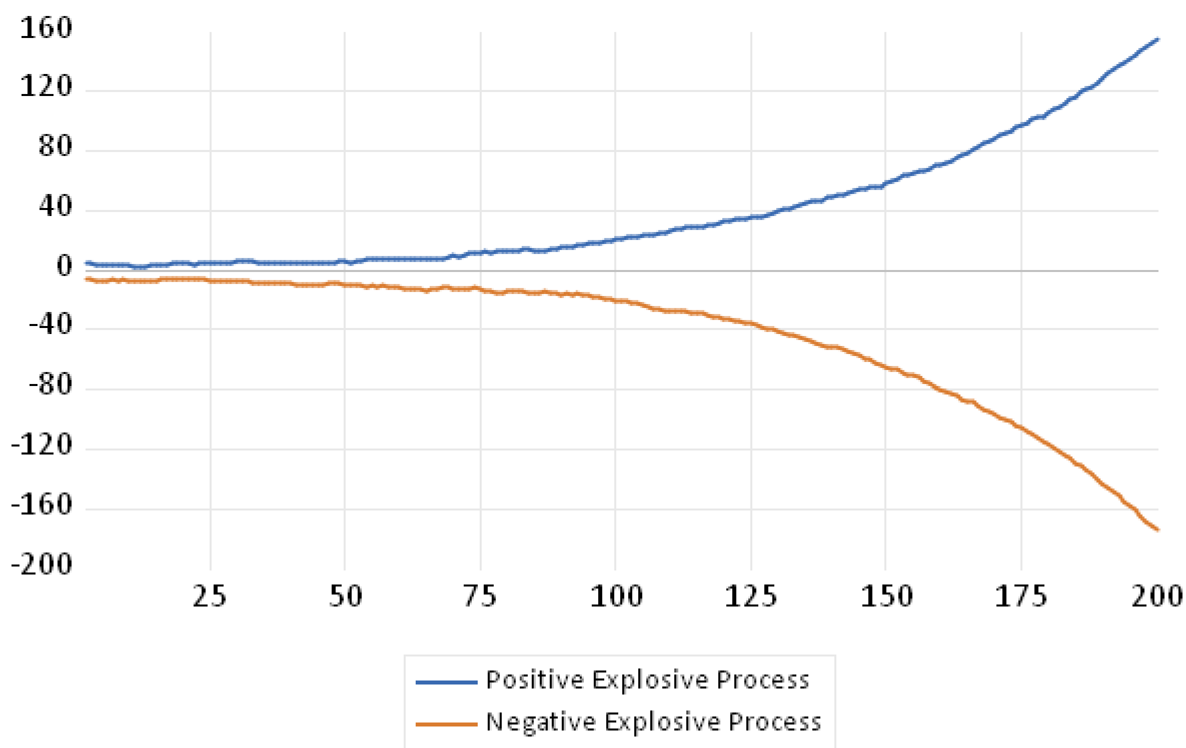


Figure 25: *Two simulated explosive processes: with positive/negative initial values*

Now as I have mentioned, the estimated value of the slope parameter in the ADF regression is invariant to the level of the points. Hence what matters is the relative position of the data points. Consider the negative explosive process from Figure 25. Because of its explosive nature, if one estimated an ADF regression on this series, the estimated value of δ would positive, and the ADF statistic would be very large[30]. If this series was shifted

[30]In fact this series was simulated with $y_t = 1.02 * y_{t-1} + \varepsilon_t$ with initial value $y_0 = -5$ and ε $N(0, 0.25)$.

up, so that all of the values were positive, the estimate of δ and the ADF statistic would be unchanged, because as I have explained the value of these statistics does not depend on the level of the points. Hence despite being shown a collapse, this series would still be identified as explosive. This is illustrated in Figure 26. Which explains why the BSADF procedure will sometimes identify the collapse periods as being explosive.

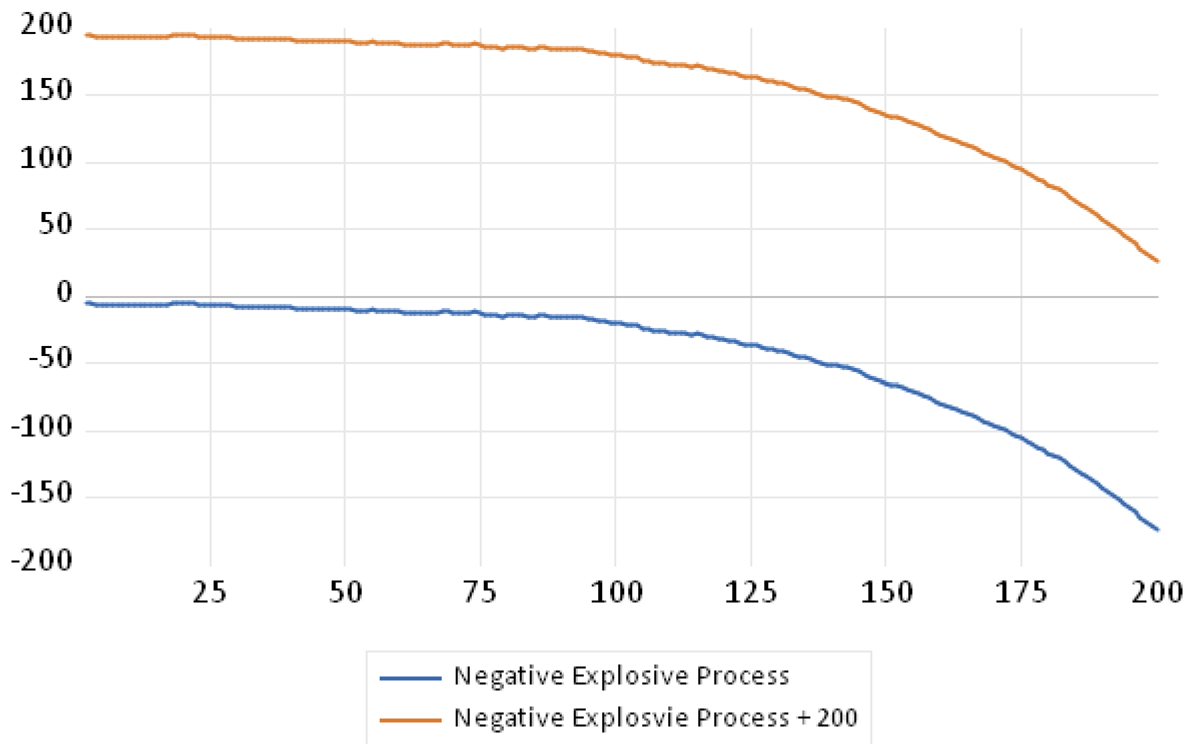


Figure 26: *Two simulated explosive processes: one negative, one identical but shifted upwards*

Deriving User Cost Equation

Starting with the condition that the current price should be equal to the discounted sum of expected future housing services and next period's housing price:

The estimated value of δ in the ADF regression (with zero lags) was 0.020584 and the ADF statistic was 26.54790

$$P_t = E_t[\frac{s_{t+1} + P_{t+1}(1 - \delta_t - \tau_{p,t})}{1 + r_{m,t} + \alpha_t}]) \tag{23}$$

Multiplying both sides by the required return, the denominator on the RHS, we get:

$$(1 + r_{m,t} + \alpha_t)P_t = E_t[s_{t+1} + P_{t+1}(1 - \delta_t - \tau_{p,t})]) \tag{24}$$

Note that $E_t(P_{t+1}) = (1 + E(\dot{P}_t))P_t$. Substituting this in, we get:

$$(1 + r_{m,t} + \alpha_t)P_t = E_t[s_{t+1} + (1 + E(\dot{P}_t))(P_t)(1 - \delta_t - \tau_{p,t})]) \tag{25}$$

Taking everything except $E_t(s_{t+1})$ over to the LHS, we get:

$$(1 + r_{m,t} + \alpha_t)P_t - (1 + E(\dot{P}_t))(1 - \delta_t - \tau_{p,t})(P_t) = E_t[s_{t+1}]) \tag{26}$$

Note that the LHS can be factorised with P_t to:

$$P_t((1 + r_{m,t} + \alpha_t) - (1 + E(\dot{P}_t))(1 - \delta_t - \tau_{p,t})) = E_t[s_{t+1}]) \tag{27}$$

Expanding out $(1 + E(\dot{P}_t))(1 - \delta_t - \tau_{p,t})$ we get that:

$$(1 + E(\dot{P}_t))(1 - \delta_t - \tau_{p,t}) = (1 - \delta_t - \tau_{p,t} + E(\dot{P}_t) - E(\dot{P}_t)\delta_t - E(\dot{P}_t)\tau_{p,t}) \tag{28}$$

Note that $E(\dot{P}_t)$, δ_t and $\tau_{p,t}$ are all very small, so $E(\dot{P}_t)\delta_t$ and $E(\dot{P}_t)\tau_{p,t})$ are both approximately equal to zero, so

$$(1 + E(\dot{P}_t))(1 - \delta_t - \tau_{p,t}) \approx (1 - \delta_t - \tau_{p,t} + E(\dot{P}_t)) \tag{29}$$

Substituting this in, we get

$$P_t((1+r_{m,t}+\alpha_t)-(1-\delta_t-\tau_{p,t}+E(\dot{P}_t))=E_t[s_{t+1}]) \tag{30}$$

This simplifies to:

$$P_t((r_{m,t}+\alpha_t)+\delta_t+\tau_{p,t}-E(\dot{P}_t))=E_t[s_{t+1}]) \tag{31}$$

If the user cost of housing is equal to the expected value of housing services, $UC_t = E_t[s_{t+1}]$, we get:

$$UC_t = P_t(r_{m,t}+\alpha_t+\delta_t+\tau_{p,t}-E(\dot{P}_t)) \tag{32}$$

Results with Different Risk Premiums

Figures 27 to 38 and Tables 6 to 9 show the results with different risk premiums used. This was done for four different risk premiums: variable using the standard CAPM and Consumption CAPM, and fixed with 2% annually and 5% (2% is often used in the literature, and 5% seems like a reasonable upper bound). Clearly the overall results are extremely similar to the results presented in Section v, where the risk premium was excluded. The main difference is that different risk premiums can shift the general *level* of the ratio. While this does not greatly affect when and where bubbles are identified, it does affect when housing becomes under or overpriced. For instance, with no risk premium, the user cost in Kensington & Chelsea & Hammersmith & Fulham becomes around zero in 2016. With *fixed* risk premiums, the ratio is still well above one at this point, despite the stark decrease. With the CAPM and CCAPM risk premiums, the ratio is slightly *negative* at this point -

because these risk premiums can be negative.

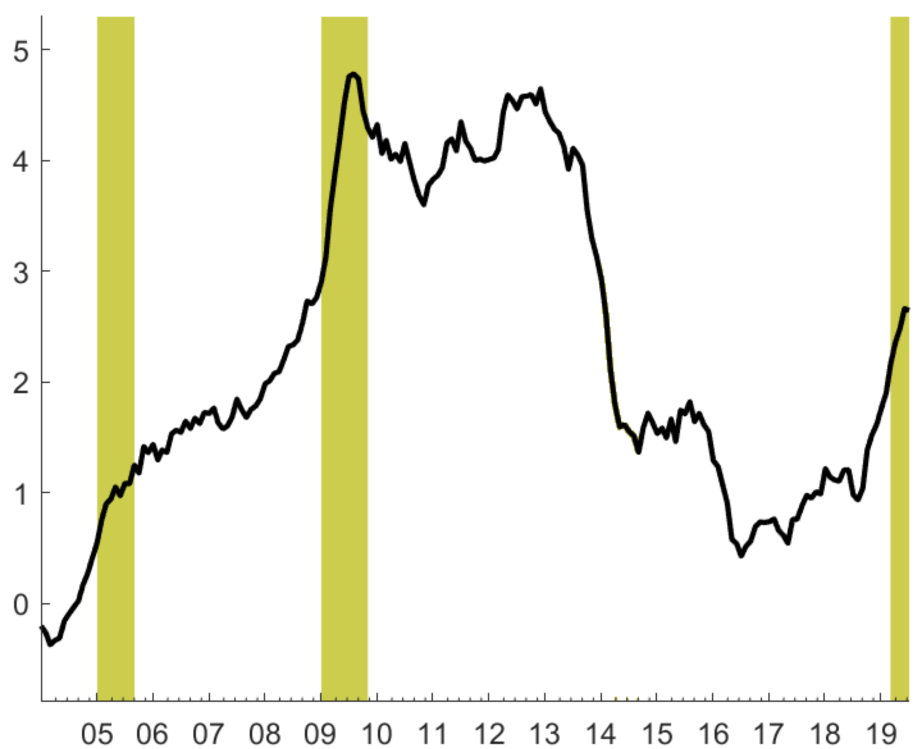
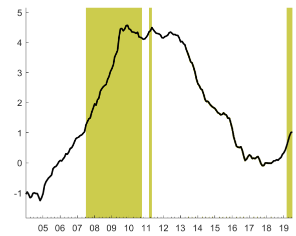
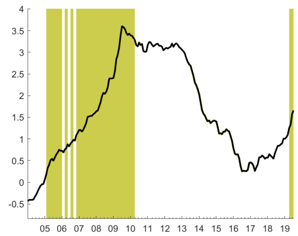


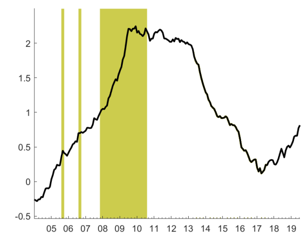
Figure 27: *User cost to rent ratio for London (with 2% fixed risk premium)*



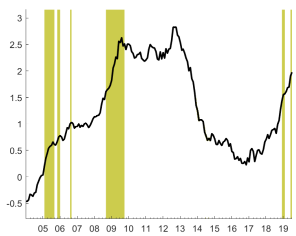
(a) *User cost to rent ratio for Barking and Dagenham and Havering*



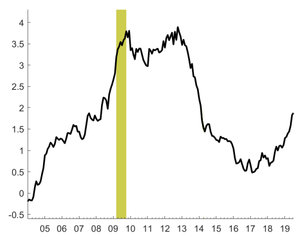
(b) *User cost to rent ratio for Barnet*



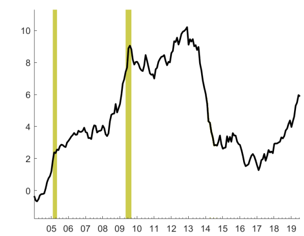
(c) *User cost to rent ratio for Bexley and Greenwich*



(d) *User cost to rent ratio for Brent*

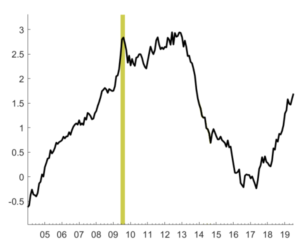


(e) *User cost to rent ratio for Bromley*

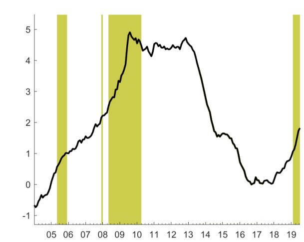


(f) *User cost to rent ratio for Camden and City of London*

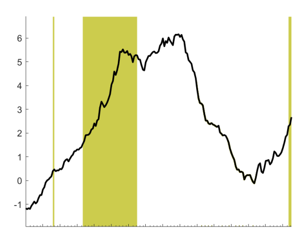
(g) *User cost to rent ratio for Croydon*



(h) *User cost to rent ratio for Ealing*

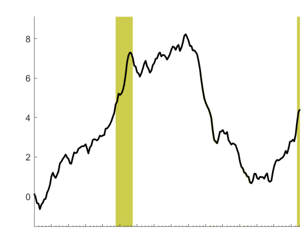


(i) *User cost to rent ratio for Enfield*



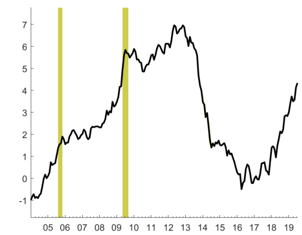
(j) *User cost to rent ratio for Hackney and Newham*

(k) *User cost to rent ratio for Haringey and Islington*

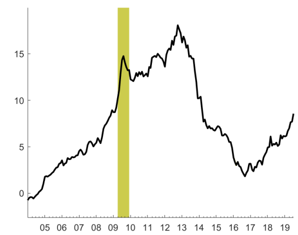


(l) *User cost to rent ratio for Harrow and Hillingdon*

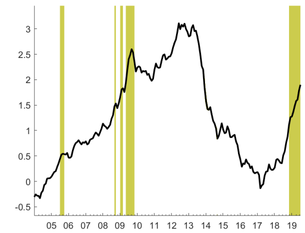
Figure 28: *User cost to rent ratio figures (with 2% fixed risk premium) (i)*



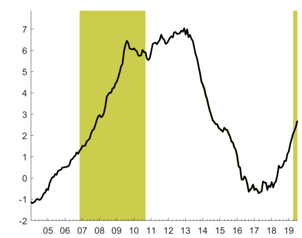
(a) *User cost to rent ratio for Hounslow and Richmond upon Thames*



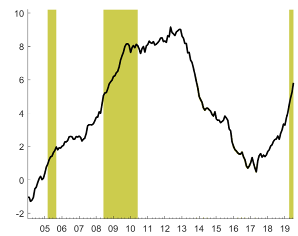
(b) *User cost to rent ratio for Kensington and Chelsea and Hammersmith and Fulham*



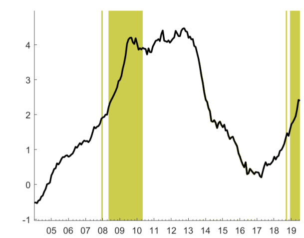
(c) *User cost to rent ratio for Lambeth*



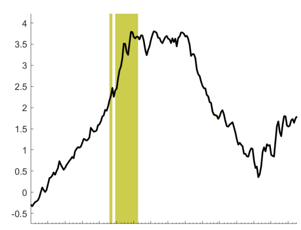
(d) *User cost to rent ratio for Lewisham and Southwark*



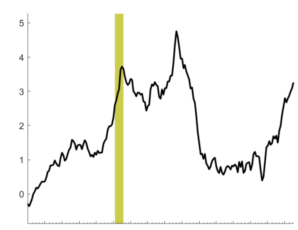
(e) *User cost to rent ratio for Merton, Kingston upon Thames and Sutton*



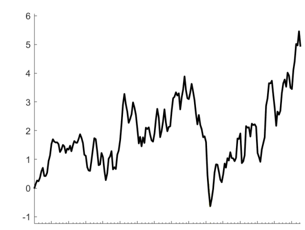
(f) *User cost to rent ratio for Redbridge and Waltham Forest*



(g) *User cost to rent ratio for Tower Hamlets*



(h) *User cost to rent ratio for Wandsworth*



(i) *User cost to rent ratio for Westminster*

Figure 29: *User cost to rent ratio figures (with 2% fixed risk premium) (ii)*

User Cost to Rent Ratio (Two percent Risk Premium)		
Region	GSADF Statistic	Critical Value
Barking and Dagenham and Havering	**5.505859**	1.580773
Barnet	**6.357824**	1.324465
Bexley and Greenwich	**5.098399**	1.396361
Brent	**3.620335**	1.425413
Bromley	**2.297279**	1.537802
Camden and City of London	**3.33251**	1.466097
Croydon	**2.643652**	1.642486
Ealing	**2.940056**	1.414438
Enfield	**5.489091**	1.608174
Hackney and Newham	**4.064544**	1.273762
Haringey and Islington	**8.15502**	1.592511
Harrow and Hillingdon	**4.18**	1.62117
Hounslow and Richmond upon Thames	**3.333394**	1.495095
Kensington and Chelsea and Hammersmith and Fulham	**4.683207**	1.558003
Lambeth	**3.607679**	1.560381
Lewisham and Southwark	**5.629748**	1.579157
Merton Kingston upon Thames and Sutton	**3.94489**	1.584251
Redbridge and Waltham Forest	**5.79146**	1.433657
Tower Hamlets	**3.838376**	1.662368
Wandsworth	**3.782551**	1.452667
Westminster	**1.90214**	1.500953
London	**6.363046**	1.513544

Table 6: *User cost to rent ratio (with 2% fixed risk premium) GSADF statistics*

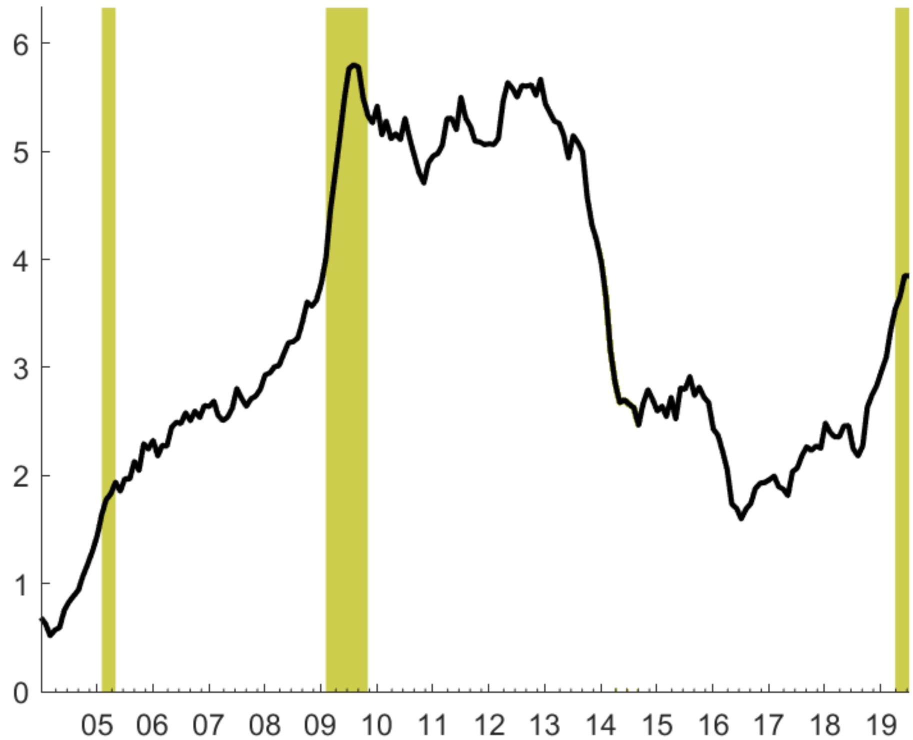
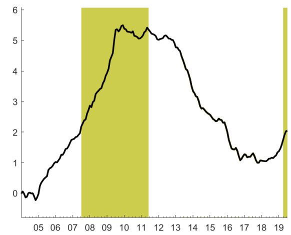
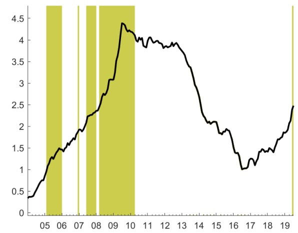


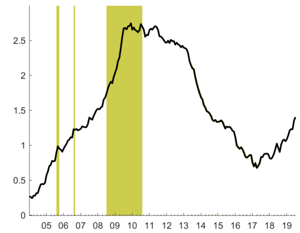
Figure 30: *User cost to rent ratio for London (with 5% fixed risk premium)*



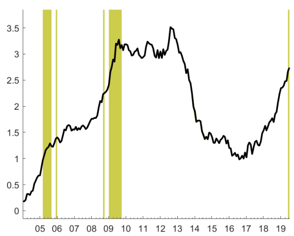
(a) *User cost to rent ratio for Barking and Dagenham and Havering*



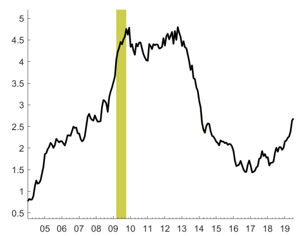
(b) *User cost to rent ratio for Barnet*



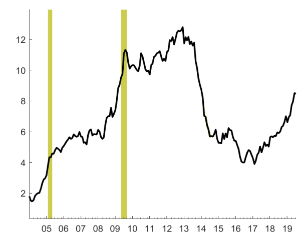
(c) *User cost to rent ratio for Bexley and Greenwich*



(d) *User cost to rent ratio for Brent*

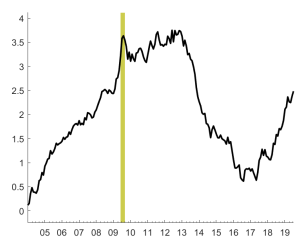


(e) *User cost to rent ratio for Bromley*

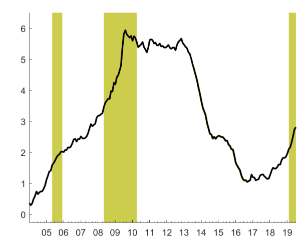


(f) *User cost to rent ratio for Camden and City of London*

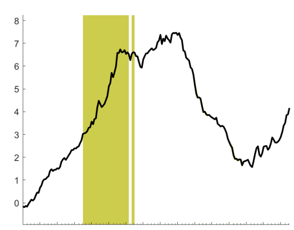
(g) *User cost to rent ratio for Croydon*



(h) *User cost to rent ratio for Ealing*

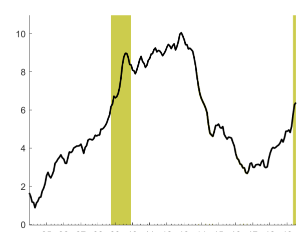


(i) *User cost to rent ratio for Enfield*



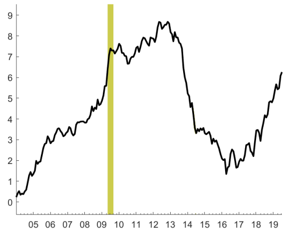
(j) *User cost to rent ratio for Hackney and Newham*

(k) *User cost to rent ratio for Haringey and Islington*

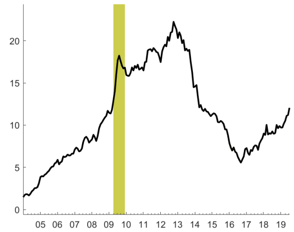


(l) *User cost to rent ratio for Harrow and Hillingdon*

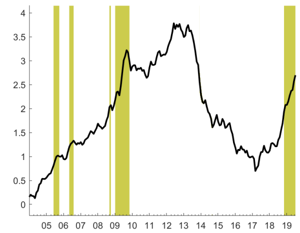
Figure 31: *User cost to rent ratio figures (with 5% fixed risk premium) (i)*



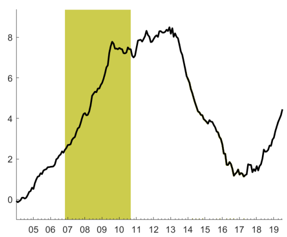
(a) *User cost to rent ratio for Hounslow and Richmond upon Thames*



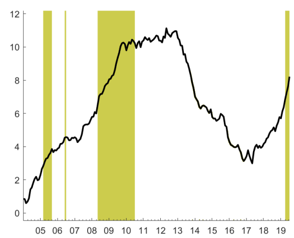
(b) *User cost to rent ratio for Kensington and Chelsea and Hammersmith and Fulham*



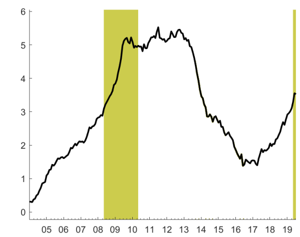
(c) *User cost to rent ratio for Lambeth*



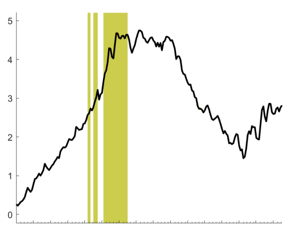
(d) *User cost to rent ratio for Lewisham and Southwark*



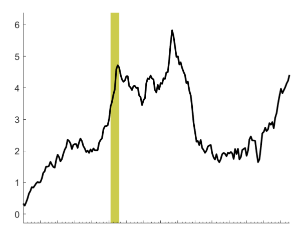
(e) *User cost to rent ratio for Merton, Kingston upon Thames and Sutton*



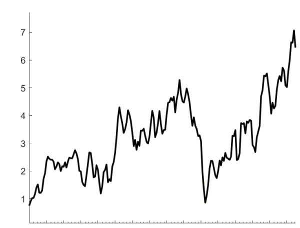
(f) *User cost to rent ratio for Redbridge and Waltham Forest*



(g) *User cost to rent ratio for Tower Hamlets*



(h) *User cost to rent ratio for Wandsworth*



(i) *User cost to rent ratio for Westminster*

Figure 32: *User cost to rent ratio figures (with 5% fixed risk premium) (ii)*

User Cost to Rent Ratio (Five percent Risk Premium)		
Region	GSADF Statistic	Critical Value
Barking and Dagenham and Havering	**5.636886638**	1.485929629
Barnet	**5.955448503**	1.549341874
Bexley and Greenwich	**5.649807303**	1.517984545
Brent	**3.592611683**	1.565091215
Bromley	**2.299838671**	1.620023815
Camden and City of London	**3.154189072**	1.563053158
Croydon	**2.711963777**	1.664253836
Ealing	**3.307846432**	1.430799173
Enfield	**5.44496822**	1.573049671
Hackney and Newham	**3.87640287**	1.599787164
Haringey and Islington	**8.125244042**	1.659214015
Harrow and Hillingdon	**4.33945146**	1.392361064
Hounslow and Richmond upon Thames	**3.208709734**	1.770106118
Kensington and Chelsea and Hammersmith and Fulham	**4.679682748**	1.806134775
Lambeth	**3.620737608**	1.382773116
Lewisham and Southwark	**5.572345063**	1.722864295
Merton Kingston upon Thames and Sutton	**3.8383163**	1.436065184
Redbridge and Waltham Forest	**5.501746516**	1.562143409
Tower Hamlets	**3.715786321**	1.573085212
Wandsworth	**4.216455591**	1.44419611
Westminster	**1.878559755**	1.486762673
London	**6.608496318**	1.613380705

Table 7: *User cost to rent ratio (with 5% fixed risk premium) GSADF statistics*

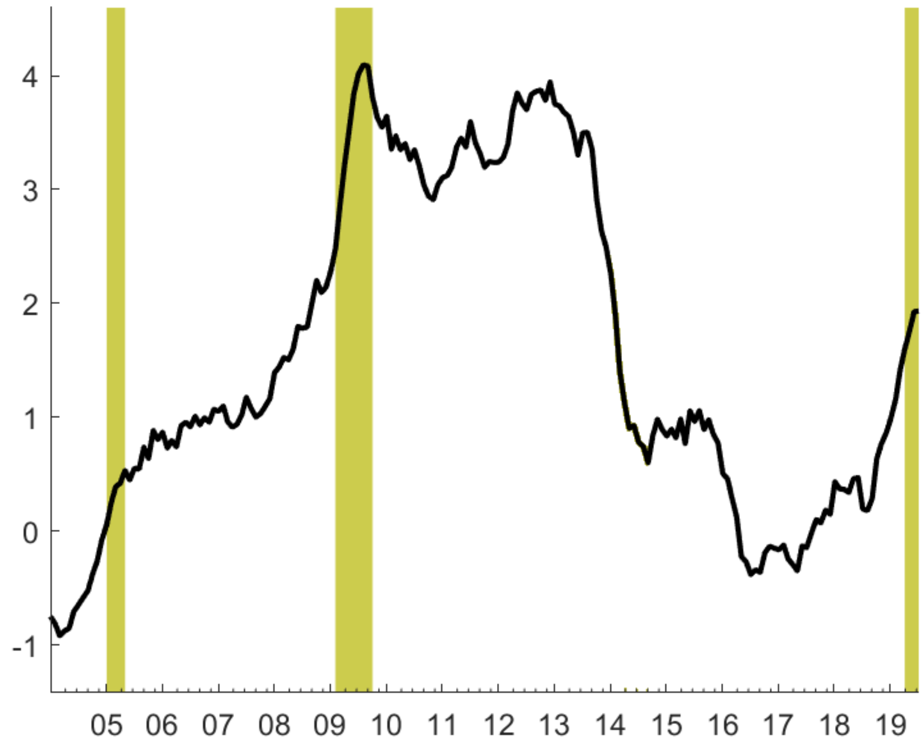
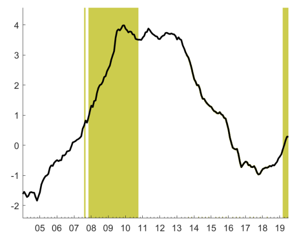
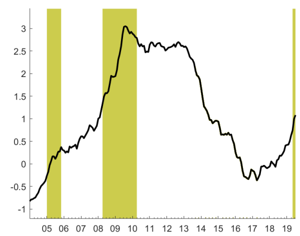


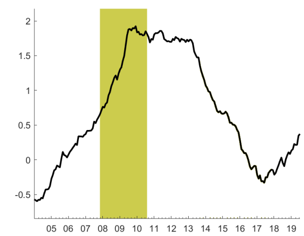
Figure 33: *User cost to rent ratio for London (with CAPM risk premium)*



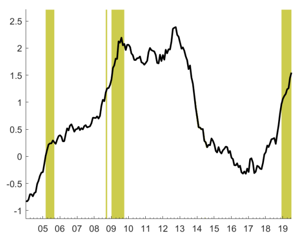
(a) *User cost to rent ratio for Barking and Dagenham and Havering*



(b) *User cost to rent ratio for Barnet*

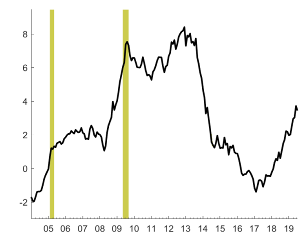


(c) *User cost to rent ratio for Bexley and Greenwich*

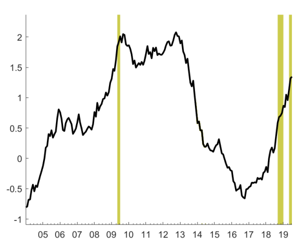


(d) *User cost to rent ratio for Brent*

(e) *User cost to rent ratio for Bromley*

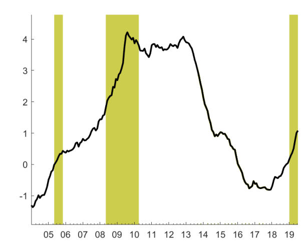


(f) *User cost to rent ratio for Camden and City of London*

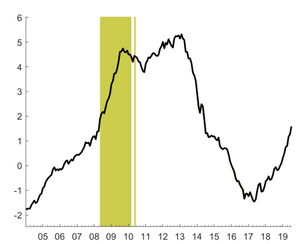


(g) *User cost to rent ratio for Croydon*

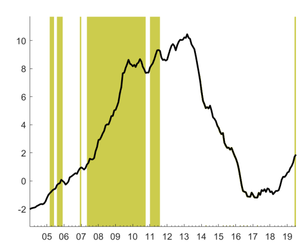
(h) *User cost to rent ratio for Ealing*



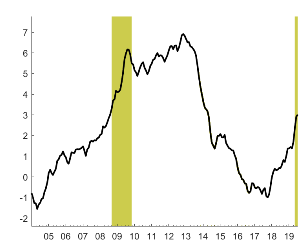
(i) *User cost to rent ratio for Enfield*



(j) *User cost to rent ratio for Hackney and Newham*

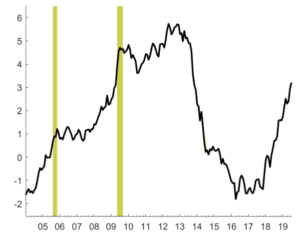


(k) *User cost to rent ratio for Haringey and Islington*

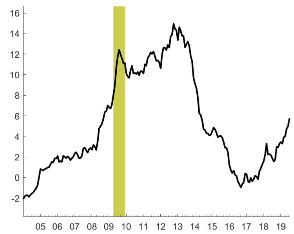


(l) *User cost to rent ratio for Harrow and Hillingdon*

Figure 34: *User cost to rent ratio figures (with CAPM risk premium) (i)*

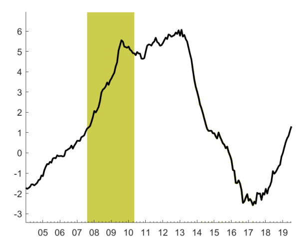


(a) *User cost to rent ratio for Hounslow and Richmond upon Thames*

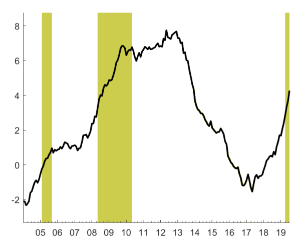


(b) *User cost to rent ratio for Kensington and Chelsea and Hammersmith and Fulham*

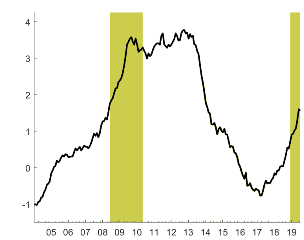
(c) *User cost to rent ratio for Lambeth*



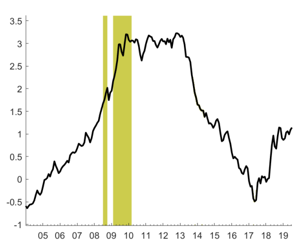
(d) *User cost to rent ratio for Lewisham and Southwark*



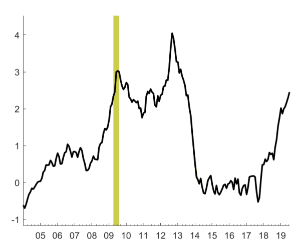
(e) *User cost to rent ratio for Merton, Kingston upon Thames and Sutton*



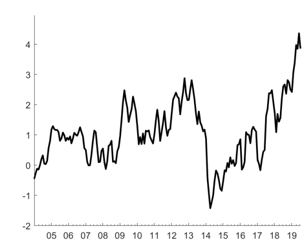
(f) *User cost to rent ratio for Redbridge and Waltham Forest*



(g) *User cost to rent ratio for Tower Hamlets*



(h) *User cost to rent ratio for Wandsworth*



(i) *User cost to rent ratio for Westminster*

Figure 35: *User cost to rent ratio figures (with CAPM risk premium) (ii)*

User Cost to Rent Ratio (CAPM Risk Premium)		
Region	GSADF Statistic	Critical Value
Barking and Dagenham and Havering	**5.536501**	1.574672
Barnet	**6.011875**	1.606858
Bexley and Greenwich	**4.449896**	1.478521
Brent	**3.359003**	1.559818
Bromley	**2.353427**	1.581152
Camden and City of London	**3.201204**	1.611601
Croydon	**2.635385**	1.558874
Ealing	**2.910371**	1.482299
Enfield	**5.343446**	1.503482
Hackney and Newham	**4.221402**	1.622076
Haringey and Islington	**6.890028**	1.675274
Harrow and Hillingdon	**3.984252**	1.34412
Hounslow and Richmond upon Thames	**3.144437**	1.441367
Kensington and Chelsea and Hammersmith and Fulham	**4.557806**	1.712248
Lambeth	**3.426896**	1.708122
Lewisham and Southwark	**5.424142**	1.618536
Merton Kingston upon Thames and Sutton	**3.540517**	1.48256
Redbridge and Waltham Forest	**4.483134**	1.351425
Tower Hamlets	**3.49898**	1.564843
Wandsworth	**3.03717**	1.646884
Westminster	**1.621279**	1.420572
London	**5.17516**	1.61398

Table 8: *User cost to rent ratio (CAPM risk premium) GSADF statistics*

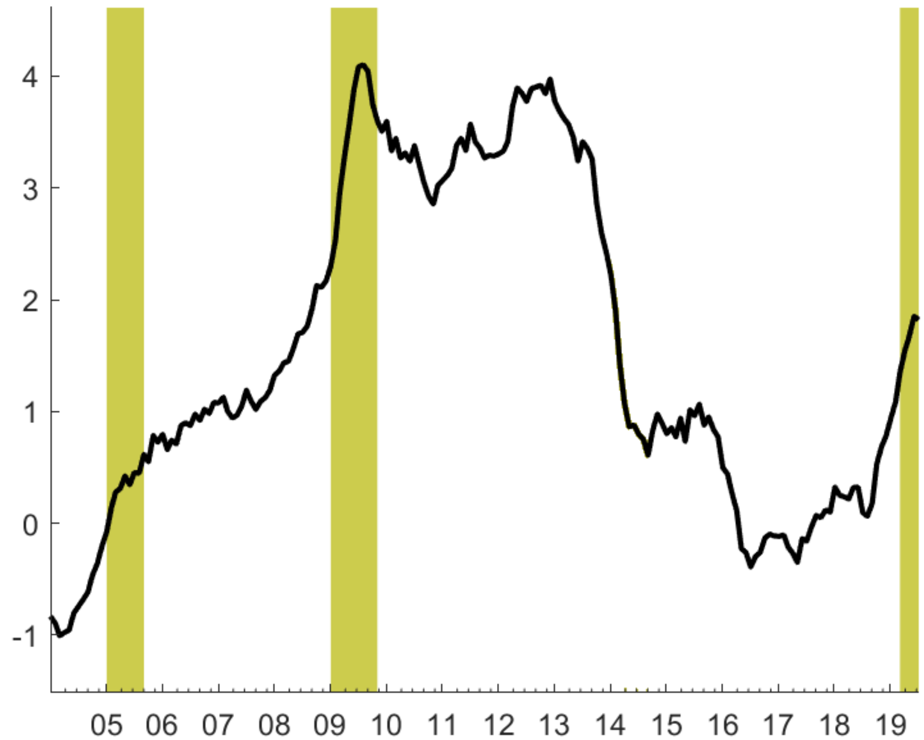
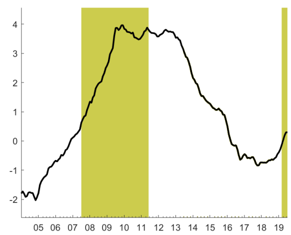
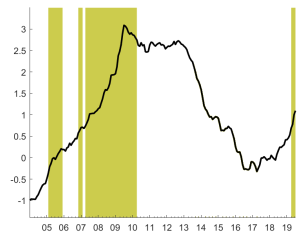


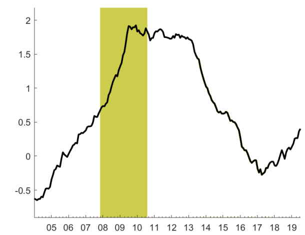
Figure 36: *User cost to rent ratio for London (with Consumption CAPM risk premium)*



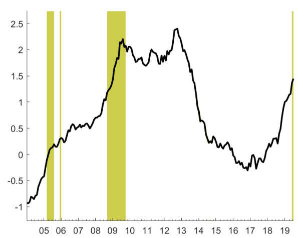
(a) *User cost to rent ratio for Barking and Dagenham and Havering*



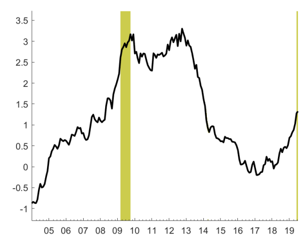
(b) *User cost to rent ratio for Barnet*



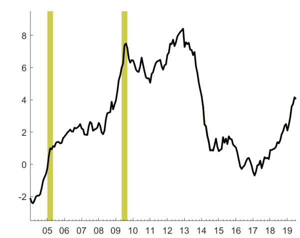
(c) *User cost to rent ratio for Bexley and Greenwich*



(d) *User cost to rent ratio for Brent*

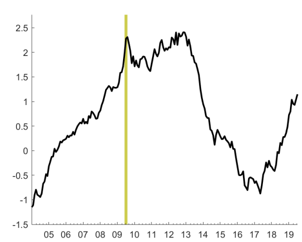


(e) *User cost to rent ratio for Bromley*

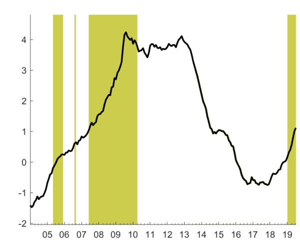


(f) *User cost to rent ratio for Camden and City of London*

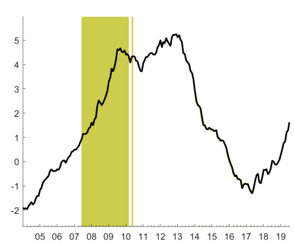
(g) *User cost to rent ratio for Croydon*



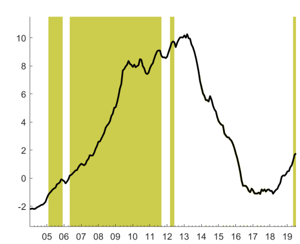
(h) *User cost to rent ratio for Ealing*



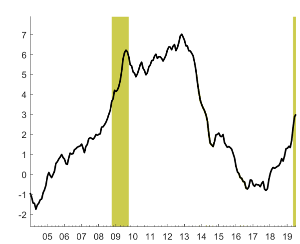
(i) *User cost to rent ratio for Enfield*



(j) *User cost to rent ratio for Hackney and Newham*

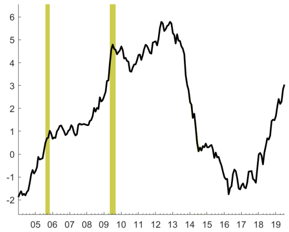


(k) *User cost to rent ratio for Haringey and Islington*

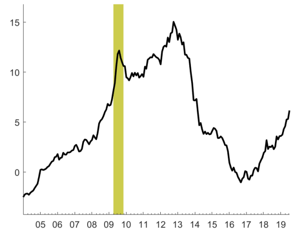


(l) *User cost to rent ratio for Harrow and Hillingdon*

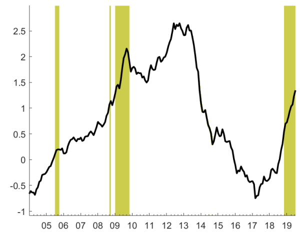
Figure 37: *User cost to rent ratio figures (with Consumption CAPM risk premium) (i)*



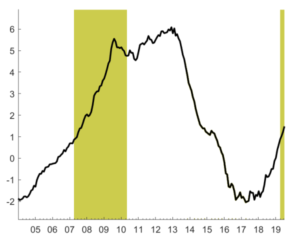
(a) *User cost to rent ratio for Hounslow and Richmond upon Thames*



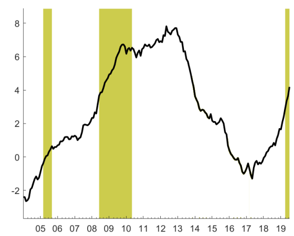
(b) *User cost to rent ratio for Kensington and Chelsea and Hammersmith and Fulham*



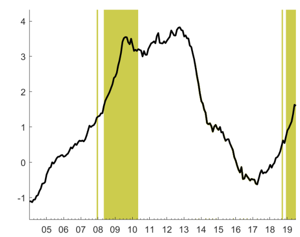
(c) *User cost to rent ratio for Lambeth*



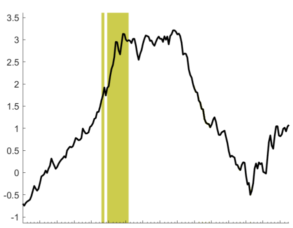
(d) *User cost to rent ratio for Lewisham and Southwark*



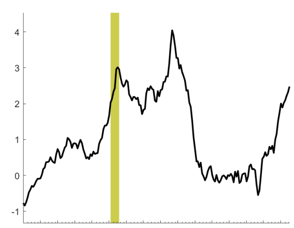
(e) *User cost to rent ratio for Merton, Kingston upon Thames and Sutton*



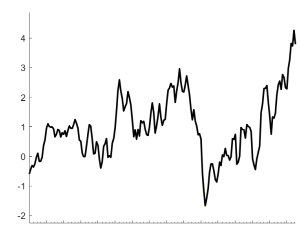
(f) *User cost to rent ratio for Redbridge and Waltham Forest*



(g) *User cost to rent ratio for Tower Hamlets*



(h) *User cost to rent ratio for Wandsworth*



(i) *User cost to rent ratio for Westminster*

Figure 38: *User cost to rent ratio figures (with Consumption CAPM risk premium) (ii)*

User Cost to Rent Ratio (CCAPM Risk Premium)		
Region	GSADF Statistic	Critical Value
Barking and Dagenham and Havering	**5.552628**	1.436864
Barnet	**6.523092**	1.5772
Bexley and Greenwich	**4.808202**	1.574663
Brent	**3.66693**	1.548823
Bromley	**2.412925**	1.531209
Camden and City of London	**3.531382**	1.638633
Croydon	**2.536758**	1.635456
Ealing	**2.778757**	1.683922
Enfield	**5.440759**	1.3964
Hackney and Newham	**4.17099**	1.4913
Haringey and Islington	**8.037728**	1.588082
Harrow and Hillingdon	**4.105698**	1.703356
Hounslow and Richmond upon Thames	**3.475954**	1.557578
Kensington and Chelsea and Hammersmith and Fulham	**4.936593**	1.752978
Lambeth	**3.541844**	1.573587
Lewisham and Southwark	**5.661591**	1.671856
Merton Kingston upon Thames and Sutton	**3.989243**	1.58396
Redbridge and Waltham Forest	**5.956497**	1.500344
Tower Hamlets	**3.871917**	1.648008
Wandsworth	**3.444368**	1.562429
Westminster	**1.894097**	1.576962
London	**6.220792**	1.496178

Table 9: *User cost to rent ratio (CCAPM risk premium) GSADF statistics*

III. Testing the Rational Housing Bubble Model across London Boroughs: Evidence of Irrational Bubbles

i. Introduction

An economic bubble is an episode of self-fulfilling price increases. The price of an asset becomes removed from its economic fundamentals, and price increases are fuelled by self-fulfilling expectations of future price increases. Many explanations of bubbles suggest that they are driven by irrational behaviour, for instance irrational expectations about future price increases. This would be where speculators have biased expectations of price increases which are removed from reality. These speculators are willing to buy the asset because they believe it will continue to increase in price indefinitely, regardless of the asset's fundamentals. Furthermore, there are theories in the behavioural finance literature which suggest that markets go through distinct periods of over- and under-confidence, and this level of confidence influences expectations about future price growth, and can lead to bubbles. Seminal works in this literature include Barberis et al. (1998), Daniel et al. (1998), Shiller (2000). Another theory of bubble formation that relies on irrational behaviour is the "greater fool" theory - where speculators are willing to buy an asset because they believe there will be someone else willing to purchase it at an even higher price (an even greater "fool"), even though they themselves may be aware that the asset's price is removed from any fundamental value[31].

[31] This idea is a very old one. The earliest use I could find of the exact phrase "greater fool theory" in this context was in New York Magazine, 31st of March 1969: "Where once men spoke solemnly of "conservation of capital," now they argue about the validity of the "greater fool theory." ... "Sure, we're paying five times

One explanation for this is that such buyers are irrationally overconfident of their own ability to trade in financial markets. Behavioural Finance theory suggests that, for instance, investors will attribute successes to their own abilities, but failures to poor luck (Hirshleifer, 2001) (Shiller, 2003).

However, there are asset pricing models that feature bubbles, but that do not assume irrationality. In these models, assets can be overpriced, in the sense that there is a component to the price (the "bubble" component) that is not justified by fundamentals. However, the overpriced component will grow such that on average a rational agent could buy the asset and expect to make an acceptable return. A model featuring bubbles and rational agents is known as a rational bubble model. A rational bubble is thus a specific kind of bubble that it would make sense for a rational agent to participate in. For a rational bubble to exist, there are specific features of the price dynamics that must hold. In this paper I test the rational bubble model across regional London housing markets using the coexplosive cointegrating procedure developed by Nielsen (2010) and Engsted and Nielsen (2012).

In the first paper of this thesis, I identified bubbles, i.e. where prices had increased in a manner not justified by fundamentals, across London boroughs during the Global Financial Crisis and the period immediately before it. Hence in this paper I limit my analysis to the period 2001 - 2011. The model of Nielsen (2010) and Engsted and Nielsen (2012) is built upon explosive price behaviour, so I first identified the precise sub-sample in this period using the PSY recursive unit root testing methodology of Phillips et al. (2011) and Phillips et al. (2015). The following structure of this paper is as follows: Section 2 reviews the literature surrounding rational bubbles. In Section 3 I describe the data. Section 4 describes

what it's worth, but somebody will soon pay us seven times what it's worth."" (Mayer, 1969, 38) the coexplosive cointegrating and PSY methodologies in more detail. In Section 5 I present the results and conclude.

ii. Literature Review

The Rational Bubble Model

The rational bubble model in its simplest form can be obtained from a standard asset pricing model with assumptions about future price increases being relaxed (Campbell et al., 1997). First consider a basic identity relating the price of an asset now to its future price plus dividends, discounted to present value by a constant expected return:

$$P_t = E_t[\frac{P_{t+1} + D_{t+1}}{1+R}] \tag{33}$$

If we iterate this forward K periods we obtain:

$$P_t = E_t[\sum_{i=1}^{K} \frac{D_{t+i}}{(1+R)^i}] + E_t[\frac{P_{t+K}}{(1+R)^K}] \tag{34}$$

Usually the term on the right is assumed to go to zero as K goes to infinity. However, this assumption can be dropped if P_t is expected to grow at a rate of R or higher, forever. If the expected growth rate of P_t is R exactly, then the term on the right has a finite value as K goes to infinity. However if the expected growth rate of P_t is faster than R forever, then the term on the right becomes infinitely large as K goes to infinity, which would result in current prices also being infinity. So in the rational bubble model, P_t is assumed to grow at rate R. Under the rational bubble model the price of an asset has two components - the discounted sum of expected future dividends, the term on the left, often called the *fundamental* component, and the bubble component, the term on the right. i.e.

$$P_t = P_{Dt} + B_t \tag{35}$$

where

$$P_{Dt} = E_t[\sum_{i=1}^{K} \frac{D_{t+i}}{(1+R)^i}] \tag{36}$$

and

$$B_t = E_t[\frac{P_{t+K}}{(1+R)^K}] \tag{37}$$

Because prices are expected to grow at rate R, so is the bubble component of prices. Hence:

$$E_t[B_{t+1}] = (1+R)B_t \tag{38}$$

Because this bubble component to prices is expected to grow at rate R, which is the required rate of return, this means that even though the asset is overpriced (in the sense that its price is not justified by its fundamentals), it would provide, on average, an acceptable return to a rational agent. Thus a rational agent would be willing to participate in such a bubble, and buy the overpriced asset.

There is much debate on whether rational bubbles can actually exist, even in *theory*. Now, obviously asset price bubbles do occur in real life. The real question is whether the existence of bubbles necessarily requires the dropping of a rational framework. Can a rational framework be established where bubbles exist? This is one of the main aims of the rational bubble literature. If one accepts that rational bubbles could exist, it then becomes an empirical question if rational bubbles are the best explanation for bubbles that are observed to occur.

Tirole (1985) is the first paper in the rational bubble literature with a full macroeconomic model (with production decisions and overlapping generations of agents making optimal investment decisions), which allows for the possibility of rational bubbles. This model relies on some specific assumptions however. The creation of a bubble in this model relies on the interest rate of the economy being lower than the growth rate, and the economy thus being dynamically inefficient. Many economists think that this does not represent most real world economies (Campbell et al., 1997). However Martin and Ventura (2018) argue that the interest rate being below the growth rate can also be produced in models that are dynamically efficient, but that have frictions. Another drawback of the Tirole (1985) model is that in it the bubble never bursts, it increases in size forever, in a deterministic fashion. Obviously this is not the case with real life bubbles, which are notable for when they burst. There are other models however, which incorporate stochastic bubbles which can burst, for instance Blanchard and Watson (1982). Blanchard and Watson (1982) was the first paper to introduce the idea of bubbles in a rational context. They propose a bubble which either grows at a rate faster than R each period, or bursts, while still having an expected growth rate of R, and also a random component each period i.e.

$$B_{t+1} = \begin{cases} \frac{1+R}{P} B_t + \varepsilon_{t+1} \text{ with probability P} \\ \\ \varepsilon_{t+1} \text{ with probability (1-P)} \end{cases} \tag{39}$$

Given that $E(\varepsilon_{t+1}) = 0$, this satisfies Equation 38.

Martin and Ventura (2018) gives a good review of macroeconomic models incorporating rational bubbles, whilst also presenting a general model with rational bubbles. In their model countries' economies go through cycles of rational bubbles, and during the bubble stage other countries invest in the bubble. They argue this explains the real world phenomenon where countries' current accounts tend to decline (i.e. due to net foreign investment) while the net wealth of a country increases (i.e. during economic booms, which might have a bubble component).

A commonly cited argument from Diba and Grossman (1988a) claims that if a rational bubble does exist, it can only exist when the asset first starts being traded. If a bubble on an asset bursts, it cannot start again. The first premise of this argument is that bubbles have to be non-negative. A negative bubble would be where a stock's market value was lower than its fundamental value. This negative bubble would grow indefinitely at rate R, and thus eventually the stock price would be expected to become negative, when the (negative) bubble component outweighs the fundamental component. If there is free disposable of assets, this should not be possible. Another firm could buy the firm with the negative bubble, transfer across assets the assets and close the bubble firm (Martin and Ventura, 2018). So it is assumed there cannot be a negative bubble with rational agents.

Furthermore, so the argument goes, from Equation 38, if there are any innovations in the bubble they must satisfy:

$$\varepsilon_{t+1} = B_{t+1} - (1+R)B_t \tag{40}$$

where

$$E_t(\varepsilon_{t+1}) = 0 \tag{41}$$

The innovations in the bubble process, ε_{t+1}, must be high enough to ensure that B_t is greater than zero at all times, given there can be no negative bubbles. Hence $\varepsilon_{t+1} \geq -(1+R)B_t$. Now suppose that the level of the bubble is currently at zero. For instance, a bubble collapsed or an asset started trading without a bubble. This means the innovations can only take values greater than or equal to zero. For any random variable, if it can only take values ≥ 0, and if it can take any value > 0 with probability greater than 0, if would have an expected value greater than zero. But the expectation of the innovations is zero (Equation 41). This means that if the bubble value is currently at zero, the *only* value any innovations can take is also zero. Because $B_{t+1} = (1+R)B_t + \varepsilon_{t+1}$, if the bubble ever has a value of zero, then both $B_t = 0$, and any innovations ε_{t+1} will also be zero, hence B_{t+1} must also be zero. So according to the argument of Diba and Grossman (1988a), if a rational bubble ever has a value of zero, then it will have a value of zero onwards, forever. So the only point that a rational bubble can start is when the asset first starts trading.

On the face of it, bubbles occurring in only new assets might not be too limiting when it comes to the stock market. For instance, the tech bubble of the late 90s and early 00s was in new internet based companies. Likewise, there is some evidence that IPOs tend to underperform the market in the long run (e.g. Loughran and Ritter (1995)), meaning that they are overpriced when first issued. This could be explained by there being a rational bubble component in the prices of newly listed stocks. Furthermore, many stocks that have gone through meteoric price rises have been listed only relatively recently e.g. Facebook and Google. So rational bubbles only occurring only in newly listed stocks isn't too much of a limitation in explaining many episodes of apparent overpricing.

However the argument from Diba and Grossman (1988a) does rule out rational bubbles explaining other forms of bubbles, such as housing. Only a very small percentage of the housing stock are new builds. Given that rational bubbles can only occur in assets that have only just started trading, if one accepts this argument then this essentially rules out rational bubbles in housing.

However, there are strong reasons to reject this argument. It crucially relies on the impossibility of negative bubbles. Diba and Grossman (1988a) states that an implication of negative bubbles would be that if there were a negative bubble, the negative bubble would inflate until it became greater than the fundamental value, at which point the price of the stock would be negative. Obviously this should not be possible with rational agents. But what if the negative bubble was relatively small? The argument is that another firm could buy the bubble firm, transfer across all the bubble firm's assets, and close the bubble firm. Of course, these events do actually happen in reality [32]. Is this possible with rational agents? There are costs associated with buying other firms and transferring across assets. It wouldn't make sense for another firm to buy another firm with a bubble value of negative £1, for instance. It might even be *impossible* for the bubble firm to be bought if the firm is sufficiently large. Suppose there was a negative bubble on Apple for instance. At the time of writing, the market capitalization of Apple is over $3 trillion dollars. There is no firm large enough that it could buy Apple and sell off its assets, if there were a negative bubble in Apple. So it seems reasonable that negative bubbles taking some small value

[32]This investment strategy is known as "asset stripping". One relatively well known investor notable for this strategy is Carl Icahn.

isn't necessarily a violation of rationality. And if a bubble can take *any* (arbitrarily small) negative value, then this nullifies the argument, because ε_{t+1} could take values other than zero (positive or negative), even if there is no bubble currently on the asset.

My view is that a rational bubble could, theoretically, exist: ultimately, all economic models are artificial simplifications of the real world. In creating an economic model, an economist will make various assumptions, and through these assumptions, the models make predictions about how the world will behave. Depending on the assumptions made, a model could predict almost anything that the model creator likes. What matters for an economic model is how realistic the assumptions are, and how the empirical data fits with the predictions the model makes. The question of whether or not rational bubbles do exist is an empirical one. We must see if the implications of rational bubble models fit the data.

Testing for Rational Bubbles

Earlier tests for rational bubbles were usually indirect i.e. the fundamental solution to the asset pricing equation would be tested i.e. if $B_t = 0$, and if the empirical data does not support this solution, then a rational bubble cannot be rejected, but if it does support the fundamental solution then the rational bubble model can be rejected. Examples of this include Diba and Grossman (1988b), who test the order of integration for prices and dividends implied by the fundamental solution, and Blanchard and Watson (1982) who test the variance bounds of stock prices compared to dividends in a similar manner to Shiller (1981), and auto-correlation of excess returns in gold prices. The basic idea with the variance bounds test is that if prices are determined by fundamentals, then the volatility of stock prices should be determined by the volatility of dividends. In reality, however, the volatility of prices is much higher than what it should be, which is evidence against the fundamental solution, which in turn means that the rational bubble hypothesis cannot be rejected. One issue with these tests, however, is that they can't distinguish between rational and irrational bubbles. If the fundamental solution is rejected, it could be that agents are merely irrational.

One way of testing the fundamental solution to the asset pricing equation is to test for a cointegrating relationship, commonly referred to as the *spread*, that is implied by the fundamental solution. Campbell and Shiller (1987) argue that with present value models, i.e. where the price of an asset is the discounted present value of future payoffs, a cointegrating relation is implied. Consider the general model

$$Y_t = \theta(1-\delta)\sum_{i=0}^{\infty}\delta^i E_t[y_{t+i}] \tag{42}$$

This means the price of the asset Y_t, is equal to the sum of expected future payoffs, y_t. δ is the discount factor, and θ is a normalizing constant. This is analogous to the fundamental solution of the asset pricing equation (Equation 34).

Define the spread as

$$S_t = Y_t - \theta y_t \tag{43}$$

Subtracting θy_t from the present value relation, Equation 42, it can be shown that the spread is equal to the expected discounted sum of changes in y_t, i.e.

$$S_t = \theta(1-\delta)\sum_{i=0}^{\infty}\Delta y_t \tag{44}$$

For more details see Appendix vi. Alternatively by iterating the present value relation forwards, it can be shown that the spread is proportional to the expected next period change, i.e.

$$S_t = \frac{\delta}{1-\delta} E_t(\Delta Y_{t+1}) \tag{45}$$

For more details see Appendix vi. If Y_t and y_t are I(1), then the spread is a cointegrating relation, as both of the RHSs of Equation 44 and Equation 45 are stationary[33]. The fundamental solution to the stock pricing model is a special case of this, where $\theta = 1/R$ and $\delta = 1/(1+R)$. Campbell and Shiller (1987) test if the spread is a cointegrating relationship for the US stock market from 1871 to 1986. They do this by first estimating a VAR system for P_t and D_t, then estimating an error correction model. From these and the sample period they estimate R, which varies significantly depending on how it was estimated. The estimated R from the VAR systems is much lower than that in the sample period, perhaps unreasonably so. They then run Engle and Granger cointegration tests, which give mixed results based on which test statistic is used. It is worth noting that they use an alternative variable, $SL_t = P_t - D_{t-1}/R$, instead of the spread (i.e. $S_t = P_t - D_t/R$ for the stock pricing model): given that the price P_t is the price at the start of the period and D_t are the dividends paid during that period, D_t is likely not available at the start of time t. $SL_t = S_t + \Delta D_t$, so if S_t and ΔD_t are stationary, so is SL_t.

Johansen and Swensen (1999) expand on this approach, and test the same US stock price data used by Campbell and Shiller. Again, they first estimate a VAR system and then cointegrating VAR systems for P_t and D_t. However here the cointegrating VAR system

[33] Cointegration is generally defined as there existing a linear combination of I(1) variables that is I(0).

is estimated repeatedly with different values of R, and an optimum value for R is then chosen by maximum likelihood. A benefit of doing this is that confidence intervals can be constructed for the value of R. With the optimum value of R, they do likelihood ratio tests on the restrictions implied by rational expectations and the fundamental solution to the stock pricing model. They find that while the model cannot be rejected, the estimated value for R (from estimating the cointegrating VAR systems) is unreasonable, similar to Campbell and Shiller (1987). Johansen and Swensen (2004) continues the method of Johansen and Swensen (1999), testing additional restrictions.

Engsted and Nielsen (2012) expand further on the Johansen and Swensen (1999, 2004) procedure, implementing the coexplosive cointegration methodology developed in Nielsen (2010). The coexplosive cointegrating model allows for cointegration to exist in a system with explosive roots. The coexplosive model operates in a similar manner to a VECM system. For instance consider a VECM model, from a VAR(k) model, such as:

$$\Delta X_t = \alpha\beta' X_{t-1} + \sum_{i=1}^{k-1} A_i \Delta X_{t-i} + \mu + \varepsilon_t \tag{46}$$

Here the vector β is the cointegrating vector, such that $\beta' X_t$ is I(0) if the variables in X_t are I(1).

The Nielsen (2010) coexplosive model has the form[34]:

$$\Delta_1 \Delta_\rho X_t = \alpha_1 \beta_1' \Delta_\rho X_{t-1} + \alpha_\rho \beta_\rho' \Delta_1 X_{t-1} + \sum_{i=1}^{k-2} \Phi_i \Delta_1 \Delta_\rho X_{t-i} + \mu + \varepsilon_t \tag{47}$$

Where $\Delta_\rho X_t \equiv X_t - \rho X_{t-1}$. Here now there are cointegrating and coexplosive vectors, β_1 and β_ρ, respectively. While cointegration is usually defined as there existing a linear

[34]Assuming that k is at least two, and that there is one cointegrating vector and one coexplosive vector.

combination of I(1) variables such that the combination is I(0), Engsted and Nielsen (2012) explain that in the coexplosive model it is rather that there are variables that have explosive and unit roots, and these variables have a linear combination that does not have a unit root.

The intuition behind the coexplosive model is that the filter Δ_1 removes a unit root, while the filter Δ_ρ removes an explosive root. X_t has both an explosive root and a unit root, so applying both filters will reduce X_t to stationarity. So the left hand side of the equation, $\Delta_1\Delta_\rho X_t$, is stationary, so the right hand side must be too. $\Delta_\rho X_{t-1}$ is unit root, as the Δ_ρ here removes the explosive root. So the cointegrating vector reduces the unit root series to stationarity. Likewise $\Delta_1 X_{t-1}$ has the unit root element filtered out but is still explosive, and the coexplosive vector reduces this to stationarity.

Theory regarding stock pricing models suggests several restrictions to the parameters of this model which can be tested in a manner similar to Johansen and Swensen (1999, 2004): Engsted and Nielsen (2012) test VAR, VECM and coexplosive models repeatedly, choosing the model with optimum parameters based on maximum likelihood, and then testing if restrictions implied by the stock pricing model can be rejected by likelihood ratio tests.

With this framework it is possible to test the cointegrating relationship implied by the fundamental solution to the stock pricing equation, in a context that allows for explosive roots from a rational bubble. As I mentioned earlier, one of the main difficulties in testing for bubbles is that usually the only way to test for bubbles is by testing the fundamental solution, and if this is rejected then rational bubbles cannot be ruled out. With the coexplosive framework, the rational bubble model can be directly tested. In Section 4, I will describe the coexplosive methodology in more detail. While the coexplosive methodology was originally done in the context of the US stock market, the same method has been applied for housing, with rents replacing dividends. For instance: Kivedal (2013), Zhang et al. (2021), Engsted et al. (2016).

iii. Data

House price data came from the UK House Price Index, which is produced by the Office for National Statistics (ONS). For England this uses sales data from HM Land Registry.

Rental data is from the imputed rental statistics produced as part of the Gross Value Added statistics which is produced by the ONS. This is calculated using private rental market data from the Valuation Office Agency (VOA) and figures for dwelling numbers from the Department for Levelling Up, Housing and Communities (DLUHC), so that the rental price produced accounts for differences across region and dwelling types. The ONS produced data is an aggregate for each region, so I used dwelling stock over time data from the DLUHC to calculate imputed rent per dwelling.

The rental data for some boroughs is merged together (for instance, Bexley and Greenwich are merged together despite being two separate boroughs). So I merged the price data together using sales figures, which is available from HM Revenue and Customs (HMRC). The Consumer Price Index (CPI), which is also produced by the ONS, was used to deflate the price data to be in real terms. All price series are in levels.

iv. Methodology

The PSY recursive unit root test

For each of the London boroughs, I applied the method of Engsted and Nielsen (2012) to test the rational bubble model across London housing markets. Full results will be presented in the next section, however for the purposes of illustrating the methodology, I will discuss the results for one region, Kensington and Chelsea and Hammersmith and Fulham, here.

I initially did these tests with the full data set used in the first paper of this thesis, which runs from July 2001 to July 2019, however for every borough the full sample is not explosive (as there are periods of significant price drops in the full sample, which generally preclude the time series being explosive). Because of this I first used the PSY recursive unit root test (developed in Phillips et al. (2011) and Phillips et al. (2015)) to identify where house prices were most explosive[35] for each borough. Essentially this involves repeatedly testing a dataset with right tailed[36] ADF tests.

The PSY procedure involves two tests, the Backwards Supremum ADF (BSADF) test and the Generalised Supremum ADF (GSADF) test. The BSADF test runs as follows: call the end of the sample being tested date r_2. Initially the sample is set to start at time 1, and the sample is tested with a right tailed unit root test[37]. The sample is tested again starting this time at time 2, and again etc. until the sample starts at time r_2 minus a minimum window length[38](denoted r_0). The BSADF test statistic is then the largest value of all the test statistics taken[39] i.e.:

$$BSADF_{r_2}(r_0) = \sup_{r_1 \in [1, r_2 - r_0]} ADF_{r_1}^{r_2} \tag{48}$$

With the GSADF test statistic, the BSADF test is repeated for different values of r_2,

[35]By which I mean where the ADF statistic is highest.

[36]With an alternative hypothesis that the data is explosive, as opposed to stationary.

[37]Lag order is chosen automatically by Akaike Information Criterion.

[38]To avoid short 'blips' in the data.

[39]Also known as the supremum, the smallest quantity in a set that is greater than or equal to all values in the set.

starting at the minimum window length to the end of the sample. The GSADF is then the largest of these statistics, i.e.:

$$GSADF_{r_2}(r_0) = \sup_{r_1 \in [1, r_2 - r_0], r_2 \in [r_0, T]} ADF_{r_1}^{r_2} \tag{49}$$

Critical values are generated using a wild bootstrap process following Phillips and Shi (2020), which allows for the possibility of heteroskedasticity.

The date range for each borough with the most explosive prices is found by finding the sub-sample of the data which has the highest ADF statistic[40]. I did this by indentifying the start and end dates corresponding to the GSADF statistic that was calculated for each borough. The explosive date range and GSADF statistic for each borough are reported in Table 25. For instance, for Kensington and Chelsea and Hammersmith and Fulham, the date range with the highest ADF statistic is 02/2005 to 09/2007. The statistic is 3.678, while the corresponding 5% critical value is 2.797, which confirms the sample is explosive[41].

[40]The reason I do this instead of using the dates that the BSADF test would indicate is because the BSADF procedure only identifies as explosive the last date in the sub-sample each time it is calculated. Using this would result in a very short sub-sample of data to use for subsequent tests.

[41]Note that these results are at odds with the results found in Chapter 2 of this thesis, where explosiveness was found in the price to income ratio from 2007 to 2008, explosiveness was not found with the price to rent ratio, and explosiveness was found in the user cost to rent ratio during 2009 to 2010. The reason that this date range was used, instead of those found in Chapter 2, is ultimately to be as charitable as possible to the coexplosive model used in this Chapter. One of the key features of the model is that the price series are explosive. However a recurring problem that occurred was that as additional restrictions were applied, the estimate of the explosiveness parameter would converge to one (as happens with Brent later in this chapter). In order to rectify this, I identified the data range with highest ADF statistic in an effort to find the section of data where prices are most explosive.

Region	Start Date	End Date	GSADF test statistic	GSADF Critical value
Barking and Dagenham and Havering	04/2004	07/2007	3.317	3.705
Barnet	11/2004	07/2007	2.857	3.347
Bexley and Greenwich	01/2012	08/2014	**3.811**	3.809
Brent	02/2004	07/2007	**3.824**	3.003
Bromley	06/2005	11/2007	3.362	3.665
Camden and City of London	02/2005	09/2007	**3.277**	2.940
Croydon	07/2005	11/2007	3.542	3.887
Ealing	12/2011	05/2014	3.132	3.151
Enfield	11/2010	04/2016	**4.121**	3.231
Hackney and Newham	12/2011	04/2014	**4.393**	3.099
Haringey and Islington	03/2010	09/2014	**4.664**	3.293
Harrow and Hillingdon	12/2011	07/2014	3.173	3.800
Hounslow and Richmond upon Thames	03/2004	08/2007	**3.942**	2.915
Kensington and Chelsea and Hammersmith and Fulham	02/2005	09/2007	**3.678**	2.797
Lambeth	03/2004	08/2007	**3.897**	2.988
Lewisham and Southwark	12/2009	08/2014	**3.425**	3.305
Merton and Kingston upon Thames and Sutton	12/2010	08/2014	**4.319**	3.308
Redbridge and Waltham Forest	08/2011	02/2014	**4.345**	3.821
Tower Hamlets	04/2005	07/2007	**5.368**	3.140
Wandsworth	03/2012	07/2014	**5.313**	3.154
Westminster	08/2009	11/2014	**3.416**	2.847
Greater London	02/2004	07/2007	**4.241**	3.903

Table 10: *PSY results for each region*

The Coexplosive Model

First, I defined $X_t = (P_t, RP_t)'$ for each of the boroughs, where P_t is the price series and RP_t is the rental price series. Following Engsted and Nielsen (2012), I estimated an unrestricted VAR(2) model:

$$X_t = A_1 X_{t-1} + A_2 X_{t-2} + \mu + \varepsilon_t \tag{50}$$

After estimating the VAR model, I ran Jarque–Bera tests to see if the residuals of the models are normally distributed. This is important because Engsted and Nielsen (2012) show that with the normality assumption, the likelihood ratio tests at the end of the methodology will have standard χ^2 distributions. The results of these are tests are reported for each borough in Table 11. For instance, for Kensington and Chelsea and Hammersmith and Fulham, the Jarque–Bera test statistic is 0.570649, with a p-value of 0.9663. So for Kensington and Chelsea and Hammersmith and Fulham, normality is not rejected, which is a prerequisite for the validity of the LR tests at the end of the Engsted and Nielsen (2012) methodology.

After this I tested for cointegration, and found the characteristic roots of the system. Testing for cointegration is important, as a key feature of the asset pricing model is that prices and rents will be cointegrated. The cointegration test employed is the Johansen cointegration test, using the trace statistic. The critical values are based on (Doornik, 1998, Table 2), who provide critical values for cointegration tests in the presence of explosive roots, following Engsted and Nielsen (2012). For Kensington and Chelsea and Hammersmith and Fulham, the trace statistic for the null hypothesis of no cointegration is 18.69153. This is

Region	Test Statistic	p-value
Bexley and Greenwich	0.349379	0.9864
Brent	**43.83956**	**0**
Camden and City of London	**13.10821**	**0.0108**
Enfield	4.093985	0.3934
Hackney and Newham	1.46467	0.8329
Haringey and Islington	5.957752	0.2023
Hounslow and Richmond upon Thames	1.071419	0.8988
Kensington and Chelsea and Hammersmith and Fulham	0.570649	0.9663
Lambeth	4.734789	0.3156
Lewisham and Southwark	0.214518	0.9946
Merton and Kingston upon Thames and Sutton	2.839052	0.5851
Redbridge and Waltham Forest	2.099061	0.7175
Tower Hamlets	4.512246	0.3411
Wandsworth	1.499345	0.8268
Westminster	2.749288	0.6006
Greater London	**41.5224**	**0**

Table 11: *Jarque–Bera test results for each region*

marginally rejected (i.e. a p-value less than 10% but greater than 5%). The trace statistic for there being at most one cointegrating equation is 5.019016, which is not rejected at any significance level. These results suggest there may be one cointegrating relation, as the asset pricing model implies. If there was cointegration present in the London region, I continued to estimate a VECM model. This model assumes that there is one explosive root and one unit root in the system, which is facilitated by reduced rank restrictions (i.e. there is one cointegrating vector). This hypothesis is denoted H_1 and the corresponding model that is estimated is denoted M_1 i.e.

$$\Delta_1 X_t = \alpha\beta' X_{t-1} + \Gamma\Delta_1 X_{t-1} + \alpha_1\zeta_1 + \varepsilon_t \tag{51}$$

I will denote the highest characteristic roots of this system ρ_1, and only continued if it was greater than one (i.e. there is an explosive root in the system, which is necessary for the coexplosive bubble model). I will also refer to the log-likelihood of the corresponding model as l_1. For Kensington and Chelsea and Hammersmith and Fulham, the highest characteristic root in the system is 1.097423, and the corresponding log-likelihood is -497.387. ρ_1 is the explosive root in the system that causes the bubble to grow at an explosive rate. Note that in this model, the rate of return, R is assumed to be fixed, and the bubble must grow at rate R, due to Equation 38. This means that the rate of return must be equal to $\rho_1 - 1$, which is equal to 9.74% for Kensington and Chelsea and Hammersmith and Fulham. Note however that the data is *monthly*. This means that a monthly rate of return of 9.74% corresponds to a yearly rate of return of above 100%, which seems absurdly high (note however, that this will vary with the different models that are estimated).

Using ρ_1, I calculated $\Delta_\rho X_t = X_t - \rho X_{t-1}$. Next, the hypothesis that rents are non-

Region	0 CEs test statistic	1 CE test statistic
Bexley and Greenwich	**26.70764****	6.373894
Brent	**20.11261***	3.194702
Camden and City of London	13.25992	3.111149
Enfield	**33.87768****	7.318319
Hackney and Newham	**24.63261****	7.259044
Haringey and Islington	**22.93455****	6.326203
Hounslow and Richmond upon Thames	12.77149	1.281625
Kensington and Chelsea and Hammersmith and Fulham	**18.69153***	5.019016
Lambeth	13.14465	3.566672
Lewisham and Southwark	12.79169	3.560799
Merton and Kingston upon Thames and Sutton	**23.15967****	**8.689231***
Redbridge and Waltham Forest	**23.24902****	4.987224
Tower Hamlets	17.79681	5.674913
Wandsworth	**24.97799****	5.720099
Westminster	**19.90542***	5.850487
Greater London	**26.38588****	3.870829
5% Critical Value	20.16	9.14
10% Critical Value	17.98	7.6

Table 12: *Cointegration test results*

explosive is tested. This is to ensure that $\alpha_\rho \beta_\rho' \Delta_1 X_{t-1}$. is stationary, with prices being explosive. This hypothesis corresponds to the restriction on the coexplosive vector such that $\beta_\rho = (0,1)'$, which is denoted H_D. The corresponding model is denoted M_{1D}:

$$\Delta_1 \Delta_\rho X_t = \alpha_1 \zeta_1 + \alpha_1 \beta_1' \Delta_\rho X_{t-1} + \alpha_\rho \Delta_1 RP_{t-1} + \varepsilon_t \tag{52}$$

A grid search over ρ is conducted for the model. This means that the model is estimated repeatedly for different values of ρ, and the value of ρ is chosen for which the corresponding model has the highest likelihood. The values of ρ that the model is estimated for must be greater than one - as otherwise the model will not be estimated due to multiple collinearity - if ρ were to equal one, then $\Delta_\rho X_t = X_t - \rho X_{t-1}$ would just be $\Delta_1 X_t$, and would appear in the model twice, making it impossible to estimate. So the lowest value for ρ in the grid search must be some value close to one - though not too close to be computationally expensive. I chose a starting value of 1.0001 as four decimal places seems to be a reasonable level of precision without being too computationally taxing - if the level of precision is greater, then the model would need to be estimated many more times. So the model is first estimated with ρ equal to 1.0001, and the log-likelihood recorded. The model is then re-estimated with ρ equal to 1.0002, and the log-likelihood recorded again, and this is repeated again for $\rho = 1.0003, 1.0004$ up until $\rho = 1.5$. I chose the final value as $\rho = 1.5$ because an explosive root of 1.5 would be essentially impossible given the apparent level of explosiveness in the data (this would imply a monthly rate of return equal to 50%). So the model is estimated repeatedly, for different values of ρ, with a range that covers all possible values of ρ (to a reasonable level of precision given computational limits). Each time the model is estimated, the log-likelihood is recorded. The value of ρ which corresponds to the model with the highest likelihood is denoted ρ_{1D} and the likelihood of the model is denoted l_{1D}. The reason I do this is because at the end of the procedure the restrictions are tested with likelihood ratio tests - if the restrictions lower the likelihood of the model too much, they are rejected. However the likelihood of the estimated model is sensitive to the value of ρ. Hence I don't want to inadvertently reject the model restrictions by estimating the model with an non-optimal level of ρ. Hence I estimate the model with the value of ρ which will give the model the highest likelihood. Figure 39 gives an example showing how likelihood varies with different values of ρ, and how there is an optimal ρ which will maximise the model likelihood. For Kensington and Chelsea and Hammersmith and Fulham, the value of ρ_{1D} is 1.0258 and the corresponding likelihood l_{1D} is -419.357. I will discuss these results in more detail in the next section.

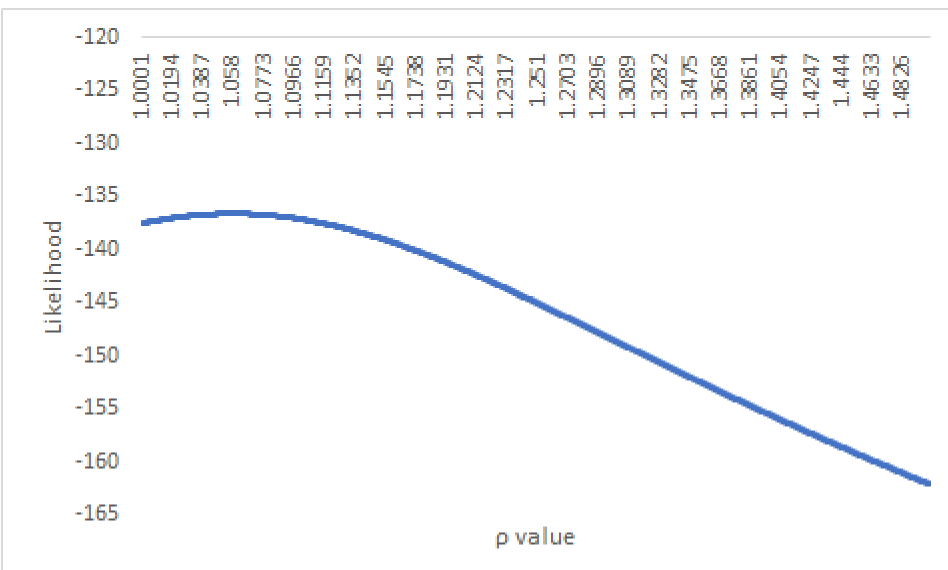


Figure 39: *An illustration of how likelihood varies with* ρ

The same model is then estimated again, but now with the restriction that $\beta_1 = (1, -1/R)'$ (i.e. the cointegrating relationship is the spread, $S_t = P_t - RP_t/R$). This is hypothesis H_S and the corresponding model is M_{1DS}. Again as with model M_{1D}, model

M_{1DS} is estimated repeatedly for different values of ρ. And similarly I denote the ρ with the maximum likelihood as ρ_{1DS} and the likelihood of the corresponding model l_{1DS}. As discussed in Section 2, this hypothesis is an important feature of the asset pricing model, and means that prices are tied to their underlying fundamentals i.e. the stream of income associated with the asset. In the case of housing, this income stream comes from rent in the case where housing is rented, and imputed rent in the case where the housing unit is owner-occupied. For Kensington and Chelsea and Hammersmith and Fulham, ρ_{1DS} is 1.0767 and l_{1DS} is -420.9037.

The final hypothesis is that $M_t \equiv P_t + RP_t - (1+R)P_{t-1}$ is a martingale difference sequence (MDS). This is another way of stating that markets are efficient. This hypothesis comes from the standard asset pricing identity relation, equating the current asset price to the price next period, plus rent, similar to Equation 33 i.e. $P_t = \frac{P_{t+1}+RP_{t+1}-M_{t+1}}{1+R}$, where M_{t+1} is a random component with expectation equal to zero. Given that markets are efficient, we would not expect these random innovations to be predictable based off of publicly available information. And this is what M_t being an MDS means - given all past information, it has an expected value of zero. This is hypothesis H_B and the corresponding model is denoted M_{1DSB}. Model M_{1DSB} is estimated from multiple equations simultaneously:

$$M_t = \varepsilon_{M,t} \tag{53}$$

$$\Delta_1 RP_t = \alpha_{1,RP}\Delta_\rho S_{t-1} + (\alpha_{\rho,RP} + \rho)\Delta_1 RP_{t-1} + \omega M_t + \varepsilon_{RP\cdot M,t} \tag{54}$$

Equation 53 follows from first pre-multiplying Equation (52) by the vector $\iota' = (1,1)$. Given that $\beta_1 = (1, -1/R)'$, this becomes:

Bexley and Greenwich	1.054119	1.017138	0.328141	-0.210307
Brent	1.086625	0.781834	0.318797 - 0.088533i	0.318797 + 0.088533i
Enfield	0.997320 - 0.047262i	0.997320 + 0.047262i	0.329581 - 0.124706i	0.329581 + 0.124706i
Hackney and Newham	0.998890 - 0.096954i	0.998890 + 0.096954i	-0.512219	-0.07262
Haringey and Islington	1.075697	0.751771 - 0.112159i	0.751771 + 0.112159i	-0.061072
Kensington and Chelsea and Hammersmith and Fulham	1.097423	0.77397	0.497542	0.14283
Merton and Kingston upon Thames and Sutton	1.037636 - 0.017627i	1.037636 + 0.017627i	0.425698	0.194977
Redbridge and Waltham Forest	1.079105	0.845768	0.410422	-0.122257
Wandsworth	1.016346	0.744161	0.105766	-0.008156
Westminster	1.032372	0.827579	0.600746	0.143546
Greater London	1.091472	0.663085 - 0.020053i	0.663085 + 0.020053i	-0.168809

Table 13: *Autoregressive roots, by region, in initial VAR system*

$$\Delta_1\Delta_\rho P_t + \Delta_1\Delta_\rho RP_t = \iota'\alpha_1\Delta_\rho S_{t-1} + \iota'\alpha_\rho\Delta_1 RP_{t-1} + \iota'\alpha_1\zeta_1 + \iota'\varepsilon_t \tag{55}$$

Given the following restrictions, this reduces to $M_t = \varepsilon_{M,t}$, where $\varepsilon_{M,t} = \iota'\varepsilon_t$:

$$H_B : \iota'\alpha_1 = -1, \iota'\alpha_\rho = -(1+R)^2/R, \zeta_1 = 0 \tag{56}$$

More details are given in Appendix vi. Equation 54 comes from the conditional equation for $\Delta_1 RP_t$ from the model Equation 52. Rearranging and defining $\varepsilon_{RP,t} = \omega\varepsilon_{RP,t} + \varepsilon_{RP\cdot M,t}$, where ω is the regression coefficient of $\varepsilon_{RP,t} = (0,1)$ on $\varepsilon_{M,t}$, gives (54).

Again model M_{1DSB} is recalculated repeatedly for different values of ρ, and the optimum value of ρ is denoted ρ_{1DSB} and the corresponding likelihood l_{1DSB}. For Kensington and Chelsea and Hammersmith and Fulham, ρ_{1DSB} is 1.0001, and l_{1DSB} is -424.8141.

Finally, the restrictions implied by each hypothesis are then be tested with likelihood ratio (LR) tests. Engsted and Nielsen (2012) show that, with some assumptions, these have standard χ^2 distributions, with degrees of freedom corresponding to the number of restrictions. What LR tests do is compare the likelihood of two models - often one is unrestricted, while the other is estimated with some restrictions on the model (alternatively one can compare two sets of restrictions). These restrictions will decrease the likelihood of the model - how much they reduce the likelihood will depend on how well the restrictions fit with the data. If the restrictions do not fit the data well, then this will heavily affect the estimated model, and the restricted model will have a much lower likelihood. This will result in the LR statistic being high and the restrictions rejected. But if the restrictions fit with the data reasonably well then the estimation of the model will not be heavily affected, if at all. The result will be the model likelihood is not reduced too much, and the resulting

LR statistic will be quite small and the restrictions will not be rejected. The results of the LR tests for Kensington and Chelsea and Hammersmith and Fulham are shown in Table 14. The LR statistic comparing model M_{1D}, which adds the restrictions implied by hypothesis H_D, that rents are non-explosive, to the unrestricted model is 2.9736. This statistic has a p-value of 8.46%, meaning that we fail to reject at the 5% significance level. The LR statistic comparing model M_{1DS}, which adds the restriction implied by hypothesis H_S to model M_{1D}, that the cointegrating relation is the spread, to the M_{1D} is 3.0934. This statistic has a p-value of 7.86%, meaning that again we fail to reject at the 5% significance level. However, the LR statistic comparing model M_{1DS}, to the unrestricted model is 6.0670, with a p-value of 4.8% which indicates a rejection at the 5% significance level. This means that the model M_{1DS} is rejected in favour of model M_1. This means that the rational bubble hypothesis is rejected for Kensington and Chelsea and Hammersmith and Fulham. I discuss this and the results for the rest of the boroughs in the next section [42].

v. Results

PSY Results

Results for the PSY procedure are reported in Table 25. Table 25 shows the GSADF statistic along with the 5% critical value for each region, along with the corresponding start date and end date of the corresponding explosive period. For most of the regions, the GSADF statistic exceeds the 5% critical value. This means that for these regions, there is evidence

[42] For completeness, I also compute the LR statistics comparing the model M_{1DSB} to the other models. For all of these statistics, the p-value is less than 5%. This means that at the 5% significance level, the restrictions implied by model M_{1DSB} would be rejected when compared to all the other models

Model	Hypothesis	Likelihood	Test Statistic	d.f.	p-value
M_1	H_1, r=1	-417.87			
M_{1D}	H_1,H_D	-419.357	LR(M_{1D}|M_1)=2.97360000000003	1	0.084633
M_{1DS}	H_1,H_D,H_S	-420.904	LR(M_{1DS}|M_{1D})=3.09339999999997	1	0.078610
			LR(M_{1DS}|M_1)=6.06700000000001	2	0.048147
M_{1DSB}	H_1,H_D,H_S,H_B	-424.814	LR(M_{1DSB}|M_{1DS})=7.82079999999996	3	0.049864
			LR(M_{1DSB}|M_{1D})=10.9141999999999	4	0.027545
			LR(M_{1DSB}|M_1)=13.8878	5	0.016338

Table 14: *LR test results, for Kensington and Chelsea & Hammersmith and Fulham*

Region	ρ_1	ρ_{1D}	ρ_{1DS}	ρ_{1DSB}
Bexley and Greenwich	1			
Brent	1.102585	1.0001	1.0001	1.0001
Enfield	1			
Hackney and Newham	1			
Haringey and Islington	1.075037	1.0076	1.0739	1.0001
Kensington and Chelsea and Hammersmith and Fulham	1.093468	1.0258	1.0767	1.0001
Merton and Kingston upon Thames and Sutton	1.078297	1.0789	1.1248	1.0002
Redbridge and Waltham Forest	1			
Wandsworth	1			
Westminster	1			
Greater London	1.076169	1.0365	1.095	1.0003

Table 15: *Values for ρ, by region*

Region	l_1	l_{1D}	l_{1DS}	l_{1DSB}
Bexley and Greenwich	-397.064			
Brent	-531.484	-539.737	-539.753	-544.859
Enfield	-794.442			
Hackney and Newham	-343.147			
Haringey and Islington	-672.975	-674.562	-676.406	-679.443
Kensington and Chelsea and Hammersmith and Fulham	-417.87	-419.357	-420.904	-424.814
Merton and Kingston upon Thames and Sutton	-515.543	-515.543	-517.549	-529.758
Redbridge and Waltham Forest	-371.303			
Wandsworth	-379.572			
Westminster	-919.273			
Greater London	-507.456	-515.209	-515.44	-517.927

Table 16: *Model log-likelihoods, by region*

of explosive price behaviour for some period of the sample. For many of the boroughs, this is a period from 2004 to 2007 - for instance in Brent, Camden & City of London, Hounslow & Richmond upon Thames, Kensington & Chelsea & Hammersmith & Fulham, Lambeth, Tower Hamlets, and Greater London. For the other regions, the period that is most explosive runs from between 2009 - 2012 until 2014 - 2016. This is the case for Bexley & Greenwich, Enfield, Hackney & Newham, Haringey & Islington, Lewisham & Southwark, Merton & Kingston upon Thames & Sutton, Redbridge and Waltham Forest, Wandsworth and Westminster.

The Coexplosive Model

The results for the Jarque–Bera tests are in Table 11. These test for normality of the errors in the model, and this is important because the LR tests used in Engsted and Nielsen (2012) assume normality of the model errors. From the test results, clearly normality can be rejected for three regions - Brent, Camden & City of London, and Greater London. This means that the assumptions used in Engsted and Nielsen (2012) are violated, and for these regions the LR test statistic at the end of the procedure would not have standard distributions, which makes inference with standard critical values invalid. At the end of this section I will discuss a method to do the Engsted and Nielsen (2012) methodology that takes account for non-normality of the errors.

Results for the initial cointegration tests, with the unrestricted model, are given in Table 12. Critical values for cointegration tests in the presence of explosive roots are based off of (Doornik, 1998, Table 2) following Engsted and Nielsen (2012). Many of the regions do not have any cointegration, and of those that do, some are marginal (only at the 10%

significance level). Results which are significant at the 5% level are bold with two asterisks. These are the regions Bexley and Greenwich, Enfield, Hackney and Newham, Haringey and Islington, Merton and Kingston upon Thames and Sutton, Redbridge and Waltham Forest, Wandsworth, Greater London. Results which are significant at the 10% level are bold with one asterisk. These are the regions Brent, Kensington and Chelsea and Hammersmith and Fulham, and Westminster. This is an interesting result, and definitely at odds with the rational bubble model for these regions. This might be interpreted as evidence of irrational bubbles for these regions, given that the prices are explosive and apparently not linked to fundamentals. Of course there are factors missing from the asset pricing model used, such as interest rates, which might explain the apparent lack of cointegration. The rational bubble model used in this paper assumes a constant interest rate, however this is not true in practice: indeed, interest rates are one of the key factors that drive house prices, through directly determining the affordability of mortgage payments. It is worth noting however that this does not apply for almost of the regions where there were explosive price increases but no cointegration, i.e. Camden and City of London, Hounslow and Richmond upon Thames, Lambeth, and Tower Hamlets. For these regions, the explosive price behaviour identified by the PSY test occurred between 2004 and 2007 - in which case interest rates actually *went up* with the UK base rate increasing from 3.7% at the start of 2004 to 5.5% by the end of 2007. In theory this should be associated with house prices falling, so changing interest rates could not justify explosive price behaviour in these periods. Likewise, changes to real incomes were modest or negative in these periods: 2.33% in Camden, -0.34% in Hounslow, 2.77% in Lambeth, -3.44% in Richmond upon Thames, and 4.29% in Tower Hamlets[43].

[43] Income data is not available for the City of London.

Given that real house price increases were at least 17% over this period for these regions, clearly increases in incomes are not enough to justify these explosive price increases.

The initial characteristic roots for each region are given in Table 13. Most of the regions have a characteristic root that is greater than 1. There are exceptions however, such as Enfield, which has highest characteristic roots of $0.99732 \pm 0.047262i$ (which is gives a modulus of 0.998439). This is a curious result, as the samples were explicitly chosen because they have explosiveness, as shown by the PSY test results. This is nonetheless possible, as the characteristic roots of a VAR system and the PSY test result are calculated in a totally different manner: the characteristic roots being the inverse eigenvalues of the coefficient matrix, while the PSY test result is ultimately based on a significantly high ADF test statistic.

If the region had cointegration present, and a characteristic root greater than one, I proceeded with the next steps, of estimating the VECM models with various restrictions. These are model M_1 which is unrestricted, model M_{1D} which has the restriction that rents are non-explosive, model M_{1DS} which is that the spread $S_t = P_t - RP_t/R$ is a cointegrating relation, and finally model M_{1DSB}, in which errors in the present value relation are a martingale difference sequence (i.e. markets are efficient).

I first estimated model M_1 for the regions with both cointegration and an explosive root in the VAR system - these are the regions Bexley and Greenwich, Brent, Enfield, Hackney & Newham, Haringey & Islington, Kensington and Chelsea and Hammersmith and Fulham, Merton & Kingston upon Thames & Sutton, Redbridge & Waltham Forest, Wandsworth, Westminster and Greater London. For many of the regions,the updated value for ρ of model M_1 is exactly one - hence there is no explosive root, which rules out the possibility of there being a bubble. In fact this makes the coexplosive system impossible to estimate, because if $\rho = 1$, then $\Delta_\rho X_t$ just becomes ΔX_t, and the model cannot be estimated due to multicollinearity. The regions where there are explosive roots in model M_1 are Brent, Haringey & Islington, Kensington and Chelsea and Hammersmith and Fulham, Merton & Kingston upon Thames & Sutton, and Greater London. I proceeded to estimate models M_{1D}, M_{1DS} and M_{1DSB} for these regions, implementing a grid search to find the optimal value for ρ (based on the value of ρ which gives the model the highest likelihood). The results for all ρ values from the VECM models, and the corresponding likelihoods, are given in Tables 15 and 16.

As I mentioned earlier, for the LR test to be valid, the residuals of the VAR model must be normally distributed, and for some of the London regions they are not. There are nine regions which have normally distributed residuals at this stage: Bexley and Greenwich, Enfield, Hackney and Newham, Haringey and Islington, Kensington and Chelsea and Hammersmith and Fulham, Merton and Kingston upon Thames and Sutton, Redbridge and Waltham Forest, Wandsworth, and Westminster. However, Bexley and Greenwich, Enfield, Hackney and Newham, Redbridge and Waltham Forest, Wandsworth, and Westminster have a highest characteristic root of one in model M_1, so the only regions for which the LR ratio tests can be applied are: Haringey and Islington, Kensington and Chelsea and Hammersmith and Fulham, and Merton and Kingston upon Thames and Sutton.

I show the results of the LR tests for these three regions in Tables 14, 18 and 19. The null hypothesis of the LR test is that the restrictions do not significantly decrease the likelihood in the restricted model vs the unrestricted (or of one restricted model compared to another restricted model). The p-value of the test statistic is the probability of observing results given the null hypothesis.

I have already discussed the results for Kensington and Chelsea and Hammersmith and Fulham in the previous section. To reiterate: The models M_{1DS} and M_{1DSB} are rejected in favour of model M_{1D}. The results for Haringey and Islington, and Merton and Kingston upon Thames and Sutton largely mirror this.

For Haringey and Islington, the LR statistic comparing model M_{1D} to the unrestricted model is 3.1740. This statistic has a p-value of 7.4819%, meaning that we fail to reject at the 5% significance level. The LR statistic comparing model M_{1DS}, to the model M_{1D} is 3.6874. This statistic has a p-value of 5.4825%, meaning that again we (marginally) fail to reject at the 5% significance level. However, the LR statistic comparing model M_{1DS} to the unrestricted model is 6.8614, with a p-value of 3.2364% which indicates a rejection at the 5% significance level. This means that the model M_{1DS} is rejected in favour of model M_1, while when comparing M_{1D} to the unrestricted model, we fail to reject. Again, for completeness, I also compute the LR statistics comparing the model M_{1DSB} to the other models. Compared to model M_{1DS}, we fail to reject model M_{1DSB} (even at the 10% significance level). However, compared to all other models, the p-value is less than 5%. This means that at the 5% significance level, the restrictions implied by model M_{1DSB} would be rejected.

For Merton and Kingston upon Thames and Sutton, the LR statistic comparing model M_{1D} to the unrestricted model is 0.0004. This statistic has a p-value of 98.40%, meaning that we fail to reject at any reasonable significance level. The LR statistic comparing model M_{1DS} to the unrestricted model is 4.0118, with a p-value of 13.4539% meaning that we fail to reject at even the 10% significance level. However, the LR statistic comparing model

M_{1DS}, to the model M_{1D} is 4.0114. This statistic has a p-value of 4.5194%, which indicates a rejection at the 5% significance level[44]. This means that the model M_{1DS} is rejected in favour of model M_{1D}. And again, I also compute the LR statistics comparing the model M_{1DSB} to the other models. Compared to all other models, the model M_{1DSB} is rejected, with a p-value of less than 5%.

So for all three of these regions, the models M_{1DS} and M_{1DSB} are rejected in favour of other models, which means that we can reject the corresponding hypotheses. The restrictions on model M_{1D} are only that there is one cointegrating vector, and that rents are non-explosive. This means that restricting the model to make rents non-explosive does not affect the model likelihood too much, and this restriction cannot be rejected. Unfortunately for the rational bubble model however, this restriction is the one most lacking in theoretical meaning for the model, and it is largely done simply to make the coexplosive model tractable, as $\alpha_\rho \beta_\rho' \Delta_1 X_{t-1}$ must be stationary. The restrictions implied by the hypotheses H_S and H_B (i.e. which generate the models M_{1DS} and M_{1DSB}) are much more economically meaningful. H_S is that the cointegrating relationship is the spread. H_B means that the housing market is efficient. Now it is worth noting that the rejection of hypothesis H_B is less problematic than the rejection of the hypothesis H_S. It has often been noted that housing markets are characterised by transaction costs, sticky prices, and other frictions that make prices slow to adjust to new information. So we would not expect the housing market to be efficient in the same way the stock market might be, even if there are no bubbles (rational or otherwise). However the rejection of the spread being a cointegrating relationship is

[44]While the statistics have a similar value, the number of restrictions is different, which means that the distribution under the null is different (A χ^2 distribution with one, instead of two, degrees of freedom)

much more catastrophic for the rational bubble model. This means that house prices are not fundamentally tied to their economic determinants, i.e. rent, as the standard asset pricing model implies. Furthermore, for all of these regions, the identified bubble period takes place during the 2004 - 2007 period. This means, as I mention earlier, this result, that the asset pricing model does not fit the data, cannot be explained away due to changing interest rates - given that interest rates went up over the period being tested for each region [45].

Accounting for non-normal data

From the results so far it seems that the hypotheses do not seem to fit the data - as the economically meaningful models M_{1DS} and M_{1DSB} are rejected for all regions which have progressed through the earlier explosiveness, normality and cointegration tests to the LR tests.

However, there were two regions for which there was explosiveness and cointegration, Brent, and Greater London (which is presumably of most interest), but with which model errors are non-normal, which makes standard inference invalid. There is a way to circumvent non-normality of the data, however. Often non-normality is caused by outliers in the data, which produce kurtosis or skewness in the distribution, which makes the distribution non-normal. If we were able to remove these outliers, then the distribution would be normal. This can be done by using impulse response dummy variables. This is done by adding in an endogenous variable that takes the value of 1 at the time of a large outlier, and zero everywhere else. This makes the error term become zero at each time the impulse response dummy takes the value of 1.

[45] A result which is corroborated by the first paper of this thesis

For both Brent and Greater London, I added one single impulse response dummy variable in January 2007 (the outlier residual occurred in the rental prices equation). I also decided to also repeat the analysis with the region Camden & City of London - while cointegration was rejected for this region, the Johansen cointegration test is also basically a likelihood ratio test, which as discussed is invalidated with non-normal residuals (given that likelihood is calculated given the assumption that the residuals are normally distributed). For Camden & City of London, I added one single impulse response dummy variable in August 2007 (again, the outlier residual occurred in the rental prices equation). Now this method is sometimes criticised for being an artificial manipulation of the data, however it is clear that there is some regularity to when the outliers occur: they all occur in the rental price equation in 2007, and two out of three occur specifically on January 2007. This might be picking up some seasonality or artifacts in the data which would usually be cleansed out in higher quality data sets. The outlier data points were from the rental dataset, which is generally of not as high quality as the price dataset. So it could be argued that by doing removing the outliers with impulse response dummies, I am in effect doing what would have been done anyway if the rental price dataset was of better quality.

Using impulse response dummies, I reran the entire procedure for the regions Brent, Camden and City of London, and Greater London. The results of this are given in Table 17.

For all these regions, the p-values of the Jarque-Bera normality tests are all at least 5%, so we cannot reject normality at the 5% significance level. There is still no evidence for cointegration in the region Camden and City of London. There is evidence of cointegration in Brent and Greater London however - particularly Greater London, where the test statistic for the null hypothesis is 32.33664, compared to a 5% critical value of 20.16. Both Brent

	Region		
Test	Brent	Camden and City of London	Greater London
Jarque–Bera test statistic	2.942859	4.770337	9.354007
Jarque–Bera p-value	0.5674	0.3117	0.0528
0 CEs test statistic	23.98739	11.9083	32.33664
0 CEs 5% Critical Value	20.16	20.16	20.16
0 CEs 10% Critical Value	17.98	17.98	17.98
1 CEs test statistic	3.838024	2.308201	4.850232
1 CEs 5% Critical Value	9.14	9.14	9.14
1 CEs 10% Critical Value	7.6	7.6	7.6
Characteristic Root 1	1.005492		1.013431
Characteristic Root 2	0.882986		0.961168
Characteristic Root 3	0.522012		0.496224
Characteristic Root 4	0.290954		-0.183775
ρ_1	1.098804		1.067543
l_1	-521.4002		-495.7344
ρ_{1D}	1.0001		1.0982
l_{1D}	-530.9627		-504.5875
ρ_{1DS}	1.0001		1.1082
l_{1DS}	-530.9837		-504.5949
ρ_{1DSB}	1.0085		1.0075
l_{1DSB}	-544.8589		-517.9268

Table 17: *Results with impulse response dummies*

and Greater London have a highest characteristic root which is greater than or equal to 1. So for both these regions, I estimated the unrestricted VECM models, and again, when estimating the VECM models, both of the highest characteristic roots remained above one.

So for both Brent and Greater London, I continued to estimate the restricted VECM models using the grid search method to obtain the models with the optimum value for ρ. The LR test results for Brent and Greater London are presented in Tables 20 and 21.

For Brent, when estimating the models M_{1D} and M_{1DS}, the coexplosive methodology starts to break down. The grid search to find the optimal ρ value is to a precision of 0.0001 - i.e. starting at 1.0001, the model is estimated and the likelihood taken, then this is repeated with ρ equal to 1.0002, etc. until ρ is equal to 1.5, which is a reasonable cut off point. For Brent, for the models M_{1D} and M_{1DS}, the highest likelihood is at $\rho = 1.0001$. What is happening here is that when these additional restrictions are added, the model is converging to non-explosiveness, i.e. a ρ equal to one. As I mentioned earlier however, these models cannot be estimated at ρ equals one exactly, hence the optimum ρ becomes as close to one as possible, which at a precision of four decimal places is 1.0001. So the model restrictions of M_{1D} and M_{1DS} apparently remove the explosiveness from the model for Brent. Nonetheless for completeness I also estimated the optimal ρ for model M_{1DSB} for Brent.

The results of the LR tests for Brent and Greater London are very similar. The models M_{1D}, M_{1DS} and M_{1DSB} are all rejected in favour of the unrestricted model, at an extremely high level of significance (i.e. a p-value of less than 0.1%). Again, this means that the rational bubble model is firmly rejected for both Greater London, and Brent.

So for most of the regions, there were explosive price increases detected in the sample, as evidenced by the GSADF tests. For these regions, the rational bubble model is rejected: either the cointegrating relationship being the spread, as the rational bubble model and general pricing theory would imply, is rejected by the likelihood ratio tests, or there is not even any cointegration whatsoever. Again, as I have mentioned, this cannot be explained away by time varying interest rates - for most of these regions the explosive period occurs during a time when interest rates were largely increasing (which the model would suggest would result in price *decreases*). This suggests that there were bubbles in these regions during this time period, and furthermore the rational bubble model does not fit the data, which excludes rational bubbles as an explanation for these price increases. Given that the fundamental solution to the asset pricing model does not apply, and the rational bubble model can be excluded, this means the best explanation of the explosive price increases observed in the data is there having been *irrational* bubbles across the London housing markets. Between 2004 -2007, house prices across Greater London, in most boroughs, went through explosive price increases that were disconnected from the economic fundamentals, in a way that cannot be explained in a rational context.

vi. Appendix

Campbell and Shiller (1987) Result I

Campbell and Shiller (1987) consider the general present value relation. Y_t is the price of the asset, which is equal to the sum of expected future payoffs, y_t. δ is the discount factor, and θ is a normalizing constant:

Model	Hypothesis	Likelihood	Test Statistic	d.f.	p-value
M_1	H_1, r=1	-672.975			
M_{1D}	H_1,H_D	-674.562	LR(M_{1D}|M_1)=3.17399999999998	1	0.074819
M_{1DS}	H_1,H_D,H_S	-676.406	LR(M_{1DS}|M_{1D})=3.68740000000003	1	0.054825
			LR(M_{1DS}|M_1)=6.8614	2	0.032364
M_{1DSB}	H_1,H_D,H_S,H_B	-679.443	LR(M_{1DSB}|M_{1DS})=6.07279999999992	3	0.108122
			LR(M_{1DSB}|M_{1D})=9.76019999999994	4	0.044667
			LR(M_{1DSB}|M_1)=12.9341999999999	5	0.024003

Table 18: *LR test results, for Haringey and Islington*

Model	Hypothesis	Likelihood	Test Statistic	d.f.	p-value
M_1	H_1, r=1	-515.543			
M_{1D}	H_1,H_D	-515.543	LR(M_{1D}|M_1)=0.000399999999899592	1	0.984043
M_{1DS}	H_1,H_D,H_S	-517.549	LR(M_{1DS}|M_{1D})=4.01140000000009	1	0.045194
			LR(M_{1DS}|M_1)=4.01179999999999	2	0.134539
M_{1DSB}	H_1,H_D,H_S,H_B	-529.758	LR(M_{1DSB}|M_{1DS})=24.4184	3	0.000020
			LR(M_{1DSB}|M_{1D})=28.4298000000001	4	0.000010
			LR(M_{1DSB}|M_1)=28.4302	5	0.000030

Table 19: *LR test results, for Merton, Kingston upon Thames and Sutton*

Model	Hypothesis	Likelihood	Test Statistic	d.f.	p-value
M_1	H_1, r=1	-521.4			
M_{1D}	H_1,H_D	-530.963	LR(M_{1D}|M_1)=19.125	1	0.000012
M_{1DS}	H_1,H_D,H_S	-530.984	LR(M_{1DS}|M_{1D})=0.0419999999999163	1	0.837620
			LR(M_{1DS}|M_1)=19.1669999999999	2	0.000069
M_{1DSB}	H_1,H_D,H_S,H_B	-544.859	LR(M_{1DSB}|M_{1DS})=27.7503999999999	3	0.000004
			LR(M_{1DSB}|M_{1D})=27.7923999999998	4	0.000014
			LR(M_{1DSB}|M_1)=46.9173999999998	5	0.000000

Table 20: *LR test results, for Brent*

Model	Hypothesis	Likelihood	Test Statistic	d.f.	p-value
M_1	H_1, r=1	-495.734			
M_{1D}	H_1,H_D	-504.588	LR(M_{1D}|M_1)=17.7062	1	0.000026
M_{1DS}	H_1,H_D,H_S	-504.595	LR(M_{1DS}|M_{1D})=0.0148000000000366	1	0.903172
			LR(M_{1DS}|M_1)=17.721	2	0.000142
M_{1DSB}	H_1,H_D,H_S,H_B	-517.927	LR(M_{1DSB}|M_{1DS})=26.6637999999999	3	0.000007
			LR(M_{1DSB}|M_{1D})=26.6786	4	0.000023
			LR(M_{1DSB}|M_1)=44.3847999999999	5	0.000000

Table 21: *LR test results, for Greater London*

$$Y_t = \theta(1-\delta)\sum_{i=0}^{\infty}\delta^i E_t[y_{t+i}] \tag{57}$$

Expanding this out, we get:

$$Y_t = \theta(1-\delta)y_t + \theta\delta(1-\delta)E_t(y_{t+1}) + \theta\delta^2(1-\delta)E_t(y_{t+2}) + \ldots \tag{58}$$

We subtract θy_t on both sides of Equation 58 to get:

$$Y_t - \theta y_t = -\theta\delta y_t + \theta\delta(1-\delta)E_t(y_{t+1}) + \theta\delta^2(1-\delta)E_t(y_{t+2}) + \ldots \tag{59}$$

Noting that $\theta\delta^K(1-\delta)E_t(y_{t+K}) = \theta\delta^K E_t(y_{t+K}) - \theta\delta^{K+1}E_t(y_{t+K})$, Equation 59 can be rewritten:

$$Y_t - \theta y_t = -\theta\delta y_t + \theta\delta E_t(y_{t+1}) - \theta\delta^2 E_t(y_{t+1}) + \theta\delta^2 E_t(y_{t+2}) - \theta\delta^3 E_t(y_{t+2}) + \ldots \tag{60}$$

Grouping the terms we get:

$$Y_t - \theta y_t = \theta\delta(E_t(y_{t+1}) - y_t) + \theta\delta^2 E_t(y_{t+2} - y_{t+1}) + \theta\delta^3 E_t(y_{t+3} - y_{t+2}) + \ldots \tag{61}$$

Furthermore, as $E_t(y_{t+1}) - y_t = E_t(\Delta y_{t+1})$, Equation 61 can be further simplified to:

$$Y_t - \theta y_t = \theta\delta E_t(\Delta y_{t+1}) + \theta\delta^2 E_t(\Delta y_{t+2}) + \theta\delta^3 E_t(\Delta y_{t+3}) + \ldots \tag{62}$$

The spread is defined as $S_t \equiv Y_t - \theta y_t$. The LHS of Equation 62 is just the spread, and the RHS can be grouped into an infinite sum to get:

$$S_t = \theta\sum_{i=1}^{\infty}\delta^i E_t(\Delta y_{t+i}) \tag{63}$$

Campbell and Shiller (1987) Result II

Starting again with the present value relation in Appendix vi:

$$Y_t = \theta(1-\delta)\sum_{i=0}^{\infty}\delta^i E_t[y_{t+i}] \tag{64}$$

By taking out the first term from the infinite sum, this can be rewritten as:

$$Y_t = \theta(1-\delta)y_t + \theta(1-\delta)\sum_{i=1}^{\infty}\delta^i E_t[y_{t+i}] \tag{65}$$

Noting that $\theta(1-\delta)\sum_{i=1}^{\infty}\delta^i E_t[y_{t+i}] = \delta\theta(1-\delta)\sum_{i=0}^{\infty}\delta^i E_t[y_{t+1+i}] = \delta E_t(Y_{t+1})$, Equation 65 becomes:

$$Y_t = \theta(1-\delta)y_t + \delta E_t(Y_{t+1}) \tag{66}$$

Noting that $Y_t = (1-\delta)Y_t + \delta Y_t$, Equation 66 becomes:

$$(1-\delta)Y_t + \delta Y_t = \theta(1-\delta)y_t + \delta E_t(Y_{t+1}) \tag{67}$$

Subtracting $(\delta Y_t + \theta(1-\delta)y_t)$ from both sides, we get:

$$(1-\delta)Y_t - \theta(1-\delta)y_t = \delta E_t(Y_{t+1}) - \delta Y_t \tag{68}$$

$S_t \equiv Y_t - \theta y_t$ and $\delta E_t(Y_{t+1}) - \delta Y_t = \delta E_t(\Delta Y_{t+1})$, so substituting these into the LHS and RHS respectively of Equation 68 we get:

$$(1-\delta)S_t = \delta E_t(\Delta Y_{t+1}) \tag{69}$$

And finally dividing through by $1-\delta$ we get the result:

$$S_t = \frac{\delta}{1-\delta} E_t(\Delta Y_{t+1}) \tag{70}$$

Deriving Model M_{1DSB} restrictions

Start with model M_{1D}:

$$\Delta_1 \Delta_\rho X_t = \alpha_1 \zeta_1 + \alpha_1 \beta_1' \Delta_\rho X_{t-1} + \alpha_\rho \Delta_1 RP_{t-1} + \varepsilon_t \tag{71}$$

First note that given $\beta_1 = (1, -1/R)'$, $\beta_1' \Delta_\rho X_t = \Delta_\rho S_t$.

Also, note that $\iota' = (1,1)$ and $X_t = (P_t, RP_t)')$, so it is the case that $\iota' \Delta_1 \Delta_\rho X_t = \Delta_1 \Delta_\rho P_t + \Delta_1 \Delta_\rho RP_t$

Pre-multiplying Equation 71 by $\iota' = (1,1)$, we get:

$$\Delta_1 \Delta_\rho (P_t + RP_t) = \iota' \alpha_1 \Delta_\rho S_{t-1} + \iota' \alpha_\rho \Delta_1 RP_{t-1} + \iota' \alpha_1 \zeta_1 + \iota' \varepsilon_t \tag{72}$$

The restrictions of H_B are as follows:

$$\iota' \alpha_1 = -1, \iota' \alpha_\rho = -(1+R)^2/R, \zeta_1 = 0 \tag{73}$$

Plugging these restrictions into Equation 72, we get:

$$\Delta_1 \Delta_\rho (P_t + RP_t) = -\Delta_\rho S_{t-1} - \frac{(1+R)^2}{R} \Delta_1 RP_{t-1} + \iota' \varepsilon_t \tag{74}$$

Note that $\Delta_1 \Delta_\rho (P_t + RP_t) = \Delta_1 (P_t + RP_t - (1+R) P_{t-1} - (1+R) RP_{t-1})$.

Substituting this into the LHS of Equation 74, we get:

$$\Delta_1(P_t + RP_t - (1+R)P_{t-1} - (1+R)RP_{t-1}) = -\Delta_\rho S_{t-1} - \frac{(1+R)^2}{R}\Delta_1 RP_{t-1} + \iota'\varepsilon_t \quad (75)$$

Adding $(1+R)\Delta_1 RP_{t-1}$ to both sides, we get:

$$\Delta_1(P_t + RP_t - (1+R)P_{t-1}) = -\Delta_\rho S_{t-1} + (1 + R - \frac{(1+R)^2}{R})\Delta_1 RP_{t-1} + \iota'\varepsilon_t \quad (76)$$

Note that $M_t \equiv P_t + RP_t - (1+R)P_{t-1}$, and $1 + R - \frac{(1+R)^2}{R} = -\frac{1+R}{R}$, so this becomes:

$$\Delta_1(M_t) = -\Delta_\rho S_{t-1} - (\frac{1+R}{R})\Delta_1 RP_{t-1} + \iota'\varepsilon_t \quad (77)$$

Expanding out $-\Delta_\rho S_{t-1}$, we get:

$$-\Delta_\rho S_{t-1} = -(P_{t-1} - \frac{1}{R}RP_{t-1} - (1+R)(P_{t-2} - \frac{1}{R}RP_{t-2})) \quad (78)$$

Note that $\frac{1}{R}RP_{t-1} = (\frac{1+R}{R} - 1)RP_{t-1}$, so this can be rearranged as:

$$-\Delta_\rho S_{t-1} = -P_{t-1} - RP_{t-1} + (1+R)P_{t-2} + \frac{1+R}{R}RP_{t-1} - \frac{1+R}{R}RP_{t-2} \quad (79)$$

Noting that $-P_{t-1} - RP_{t-1} + (1+R)P_{t-2} = -M_{t-1}$ and $\frac{1+R}{R}RP_{t-1} - \frac{1+R}{R}RP_{t-2} = \frac{1+R}{R}\Delta_1 RP_{t-1}$, this can be further simplified to:

$$-\Delta_\rho S_{t-1} = -M_{t-1} + \frac{1+R}{R}\Delta_1 RP_{t-1} \quad (80)$$

Plugging this into the RHS of Equation 77, we get:

$$\Delta_1 M_t = -M_{t-1} + \frac{1+R}{R}\Delta_1 RP_{t-1} - (\frac{1+R}{R})\Delta_1 RP_{t-1} + \iota'\varepsilon_t \quad (81)$$

Adding M_{t-1} to both sides, cancelling out the $\Delta_1 RP_{t-1}$ terms, and denoting $\varepsilon_{M,t} = \iota' \varepsilon_t$ we get: $M_t = \varepsilon_{M,t}$

IV. Testing London Housing Markets for Spatial Price Diffusion and Bubble Contagion

i. Introduction

The ripple effect is an empirical phenomenon observed in housing markets, where a price change in one region will be followed by price changes in surrounding regions. The literature focuses on London being the epicentre of these price movements in the UK, hence price movements first effect London, from where the price changes 'ripple out' into the surrounding area. There are different theoretical explanations for this effect. One explanation is that information is transmitted spatially: price changes occur as information flows through the informational network which spans across the country. For instance, London is the most globally interconnected part of the UK. Perhaps, this is the part of the country where buyers and sellers in the housing market are most up to date on global economic events. Economic news, that will affect both the global economy and the UK, comes first into London, where prices adjust. However more remote areas in the UK will incorporate changes less rapidly, and prices will adjust in these regions after the adjustment has occurred in London.

It could be argued that this makes sense in the context of how housing markets operate. Housing markets are characterised by *frictions*: transaction costs, search and matching costs, and informational asymmetries. There are many transaction fees e.g. real estate agent fees, conveyancer fees, taxes etc., however much of the frictions come from the simple fact that no two houses are exactly the same. Every house is unique, which means that significant time and effort has to be spent by a prospective buyer to search for the right house. The seller of a house will have much more information about the property than any prospective buyer, and they can be selective in disclosing information about a property. However, the prospective seller lacks useful information about the demand side of the market. This means that property transactions take a significant amount of time compared to transactions in other markets, which in turn means there will be a large lag between the conditions past houses were sold in, and current economic conditions. The initial reservation price of the seller will be contingent on their expectations about the wider economy. The key premise then, for this explanation of the ripple effect, would be that these expectations might vary across the country.

Another phenomenon that is notable in housing markets is the rapid price appreciation that has taken place in recent years. At the time of writing, house prices have appreciated 250% in the UK in the last twenty five years - in London, the figure is 347%[46]. Many commentators are calling what has happened a bubble. A bubble is a period of price increases fueled by self-fulfilling expectations: prices are increasing, and buyers think that prices are going to keep increasing, so when they buy because of this, this introduces demand which puts upward pressure on the price, which increases. The expectations of changing house prices are key to house price bubbles. A relatively recent idea in the literature is of 'coexplosivity' in bubbles. In the way that Evripidou et al. (2022) use this term, coexplosivity is where there are two explosive series, which are nonetheless tied together, in a similar manner to cointegration. With bubbles, coexplosivity would mean that a bubble in one market would transmit to other markets. Evripidou et al. (2022) use this

[46]From the UK Housing Price Index, July 1999 to July 2024.

in the context of the precious metals market: presumably there is some overlap in buyers of gold and buyers of silver etc., so perhaps market exuberance could spread from one market to another. I suggest here that perhaps this could happen spatially across housing markets. If this did occur, this would support the aforementioned narrative explaining the ripple effect: that price changes occur as expectations change across parts of the country at different times.

There are two main elements to this paper. First, I use a spatio-temporal model following Holly et al. (2011) to analyse the ripple effect in local housing markets in the Greater London area, at the local level. In this model, there is a dominant region, where house prices change first, which leads the following regions. The model also allows for price changes to come from a region's immediate neighbours. I also use a coexplosive framework following Evripidou et al. (2022). In this framework, if a linear combination of explosive series is stationary, they are said to be "coexplosive". I.e. there is some fundamental relationship that links the two explosive series together. The structure of the rest of this paper is as follows: in Section ii, I review the literature of the ripple effect, spatial models, and coexplosivity. In Section iii, I describe the data used. In Section iv, I describe the methodology of the spatio-temporal and coexplosive models. In Section v I describe the results, and Section vi concludes.

ii. Literature Review

Spatial Models

Before going into the theory and background of the ripple effect, I think it is first prudent to give a brief overview of the kind of models that are used in modelling spatial effects.

Elhorst et al. (2014) provides a good overview of the spatial models that are common in the literature associated with spatial price effects. So in my following summary of spatial econometrics, I largely follow Elhorst et al. (2014).

Spatial Econometrics is the field that deals with, in general, models in which the dependent variable can be affected by lagged dependent variables, the independent variables from a different lagged dependent variable, or where the error terms in the model are affected by (lagged) values of the error terms. Unlike in time series econometric however, these lags are across space instead of time. These spatial models naturally lend themselves to situations where one geographic region can be affected by variables in another region that is adjacent to it[47]. For instance, suppose one was modelling prices for some good across regions. The price, the dependent variable, would be dependent on different supply and demand factors in that region (the independent variables). However it might also be dependent on the price in different regions - e.g. if the price of the good becomes particularly cheap in another region, it might lower the price in the original region. So the dependent variable depends also on the dependent variable lagged across space. It might also depend on demand and supply factors in other regions. A huge uptick in demand in another region might affect prices in the original region (e.g. buyers might be willing to cross the border and buy the good in a different region). Hence the dependent variable is also dependent on spatial lagged independent variables. Finally a shock to prices in one region might simultaneously cause a shock to prices in an adjacent region. Hence some models include spatial interaction affects between the error terms.

[47]Note however the spatial dimension doesn't need to be literally in space, it could be across different markets or products, for instance.

One major point of difference from time series econometrics is that the effects of these lags can go both ways, unlike across time in which case the effect of lags only goes one way. The past can affect the future but the future cannot affect the past. However a region that is affected by an adjacent region will often also simultaneously affect that region. For this reason spatial effects usually use a spatial matrix, often denoted W. Often the values in the spatial matrix are either 1 or 0, based on whether the other region is geographically adjacent to the first or not. This spatial matrix is usually taken *a priori*, however sometimes practitioners use different spatial matrices for robustness tests. Alternatively the model could be estimated with Bayesian methods via Markov Chain Monte Carlo (MCMC), however this will not be explored here.

A general spatial model, with spatially lagged dependent variables, spatially lagged independent variables, and spatial interaction effects among error terms, in a simple cross-sectional setting, can be given by general nesting spatial model:

$$\begin{aligned} Y &= \delta WY + \alpha \iota_N + X\beta + WX\theta + u \\ u &= \lambda Wu + \varepsilon \end{aligned} \tag{82}$$

Where Y is an $N \times 1$ vector of N observations of the dependent variable, X is an $N \times K$ matrix of the independent variables, ι_N is an $N \times 1$ vector of ones, α is the constant parameter to be estimated, β is a $K \times 1$ vector of parameters to be estimated. u and ε are $N \times 1$ vectors of error terms. δ, the spatial autoregressive coefficient, and λ, the spatial autocorrelation coefficient, are parameters to be estimated. W is the spatial weights matrix. θ is a $K \times 1$ vector of parameters to be estimated.

This general model can be transformed into more specific models with restrictions on the parameters. For instance, with $\theta = 0$ and $\lambda = 0$ we get the Spatial Lag Model, also known as the Spatial Autoregressive Model (SAR):

$$Y = \delta WY + \alpha \iota_N + X\beta + \varepsilon \tag{83}$$

With the restriction $\lambda = 0$, this model becomes the Spatial Durbin Model:

$$Y = \delta WY + \alpha \iota_N + X\beta + WX\theta + \varepsilon \tag{84}$$

With $\theta = 0$ and $\delta = 0$ we get the Spatial Error Model:

$$\begin{aligned} Y &= \alpha \iota_N + X\beta + u \\ u &= \lambda Wu + \varepsilon \end{aligned} \tag{85}$$

These models can be extended to panel datasets by putting in a time subscript for the variables and error terms, and adding in time and spatial specific effects. These models can also be extended into dynamic spatial panel models, which include temporally lagged variables and/or variables that are lagged both temporally and spatially. These extensions complicate the estimation of the models, however.

The Ripple Effect in Housing Markets

Meen (1996) and Meen (1999) are some earlier papers analysing the so-called "ripple effect" in regional housing markets in the UK, though it is noted by both papers that by this point the empirical existence of the ripple effect as a statistical phenomenon (via cointegration) was well documented, despite theoretical explanations being lacking. Meen (1996) and Meen (1999) are nonetheless early in addressing potential theoretical explanations for the ripple effect. Meen (1996) is also an earlier paper that differentiates the different kind of econometric relationships that could be described as a ripple effect. Spatial dependence broadly means price changes in one region will affect price changes in another region - spatial convergence is where there is a long run relationship tying the prices in two regions together (i.e. there is cointegration).

Meen (1996) analyses spatial aggregation, convergence and dependence across and between regional housing markets in England. They analyse both housing starts and house prices. They find evidence of long run convergence in house prices (via cointegration tests), despite significant short-term divergence. Likewise, using a Spatial Lag Model, they find that there is unidirectional spatial dependence from the South-East to the rest of England. They also find similar results for housing starts, though the relationships are more complicated. They also discuss what it means for a national housing market to exist, in terms of spatial coefficient heterogeneity. They posit that if the differences in coefficients (in user cost etc. equations) for sub-regions are sufficiently small such that over time and over the larger region they cancel out, then it can be said that a national market exists (and thus aggregation is valid). This result is at the broader level of English sub-regions. Meen (1999) discusses economic explanations for the ripple effect. In general, there are four main explanations for the ripple effect. One explanation is what Meen (1999) refer to as "spatial arbitrage", which is essentially the process I described in the introduction. To paraphrase Meen (1999): if housing markets were totally efficient, house prices in one region could not be used to predict house prices in another. In practice however, search costs might create a diffusion process similar to the ripple effect, as information is transmitted contiguously across spatial grounds. Another explanation for the ripple effect is migration.

People will move from areas with higher priced housing to areas with lower priced housing. One issue with this however is that internal migration flows in the UK are fairly weak, and furthermore house price differentials are not a major motivator for internal migration. Another explanation that uses elements of both these is the "equity transfer" hypothesis. This is where price shocks first affect more affluent regions, where house prices are higher, and increased equity of the homeowners in these areas allows them to finance new house purchases in different regions. This could be through migration, or could be done without physically moving e.g. buying a second home to rent out. Finally the most basic explanation is that there are underlying spatial trends in the economic determinants of house prices. In this case, house price differentials are just a product of economic differentials, which have a spatial component e.g. suppose income shocks come first to SE England and then radiate to the North. This would cause house prices to rise first in the SE, followed by a rise in the North, without house prices in SE England directly affecting those in the North. Meen (1996) argues that the ripple effect could be explained by spatial coefficient heterogeneity. Using an error correction model, they find that the ripple effect can be explained at least in part due to different regions having different responses to changes in the underlying economic fundamentals. For example, the SE has a stronger response to interest rate shocks than the North (explained by the SE having higher levels of leverage). So if there is an interest rate shock, then house prices will adjust more rapidly in the SE than the North, but as the effect of the shock dissipates over time both will return to the baseline level. Hence they argue that the ripple effect can be explained by different reactions to economic shocks that take place within a region, as opposed to interactions that take place between regions e.g. migration.

Cameron et al. (2006) is another, relatively early, paper that models housing markets with a spatial model. Cameron et al. (2006) estimate an inverted demand model (where prices are a function of demand factors, mainly housing stock, incomes, and the user cost of housing), which includes spatial lags in price differentials (and spatial coefficient heterogeneity). They find significant interaction effects, with London as the leading region (though curiously one of the main conclusions of Cameron et al. (2006) is that there was *not* a bubble in UK housing at this time, even though UK house prices would implode, along with global real estate markets, shortly after publication[48]).

A very popular paper that analyses the temporal-spatial relationships of regional housing markets in the UK is Holly et al. (2011). Holly et al. (2011) use what they refer to as a spatio-temporal price diffusion model, and show how shocks to London housing prices propagate through the rest of the UK. Their model operates in a similar fashion to a spatial vector autoregressive system. London is treated separately from the other regions as the "dominant region", with the other UK regions having error-correction terms for the effect of a shock to London prices. They show that London house prices are long-run causal for house prices in the rest of the UK. Additionally, they repeat the analysis with New York house prices having a similar relationship to London prices as London has to the rest of the UK. They show that shocks to New York have an effect on London house prices (but no effect on the rest of the UK). They argue that this is explained by global financial shocks (e.g. hitting New York first, globally) first effecting the UK through London, which then

[48]In defence of Cameron et al. (2006) however, this conclusion was largely predicated upon comparing model predicted to actual house prices in 2003 - and finding that model predicted house prices were not much higher than the model predicted. However house prices peaked in 2007, with much of the price appreciation occurring in the years between 2003 and 2007.

propagates to the rest of the world.

Some more recent papers that have analysed the ripple effect in the UK housing market include Tsai (2014) and Cook and Watson (2016). Tsai (2014) investigate the ripple effect in both house prices and housing transaction volumes in the UK, at the regional level. With panel unit root tests, they find that the ratios of regional house prices to national house prices (and regional transaction volumes to national transaction volumes) are stationary. They argue that this is evidence of long-run convergence in transaction volumes and prices, and is therefore indicative of there being an equilibrium relationship and ripple effect. Cook and Watson (2016) analyse the ripple effect in the *changes* in house prices, as opposed to price *levels*, at the regional level of the UK. Using a directional forecast model, they find that the closer a region is to London, the greater the co-movement in changes in house prices, particularly in the south-east of the UK.

Similar tests have been done in housing markets across the world, with varying results. Many of these tests are relatively basic, testing either stationarity of price ratios, or for cointegration between price series. Bangura and Lee (2020) analyse housing sub-markets across Greater Sydney, with a methodology following Meen (1999). They find that the ratio of the lower price sub-market price to higher price sub-market is stationary over time, so the two prices converge together, in a ripple effect. They do Granger causality tests in a VECM context, and find that price changes in the lower price sub-market Granger-cause price changes in the higher priced sub-market. In addition to this, they find that prices in the lower price sub-market are more sensitive to fundamentals than prices in the higher price sub-market. They argue that this is evidence of an equity transfer process, where changes in economic fundamentals affect the home prices of the lower priced sub-market first, which is then transmitted to the higher price sub-market. For instance with a positive shock, this would increase home valuations in the lower priced sub-market first, the homeowners of which can then more easily trade up for a house in the higher priced sub-market. Balcilar et al. (2013) test for ripple effects in metropolitan areas in South Africa. They use Bayesian and non-linear unit root tests[49] on the ratios of metropolitan area house prices to national house prices, with the interpretation that stationarity implies convergence. They find evidence of stationarity, and therefore ripple effects, for all of the regions. Lean and Smyth (2013) test for ripple effects at the regional level in Malaysia. They perform Lagrange Multiplier unit root tests with structural breaks on the ratio of regional house prices to national house prices in the regions they test. They find evidence of stationarity for some of the regions, and evidence of regional clustering. Chiang (2014) test for ripple effects between first tier cities in China. Using a VECM model, they find evidence of long run relationships between almost all of the cities, and using Toda–Yamamoto causality tests they find that house prices changes in Beijing cause the price movements in other cities. Chen et al. (2011) find evidence of cointegration between house prices across cities in Taiwan, and argue that this is evidence of the ripple effect. Chiang and Tsai (2016) test for ripples effects in major metropolitan areas and associated major cities in different regions of the US. Using a VAR model and granger causality tests, they find that shocks to the major cities will transmit to the surrounding metropolitan area, and shocks to the Los Angeles also transmit to other areas of the United States. Using an Autoregressive Distributional-Lag (ARDL) model similar to the one used by Holly et al. (2011), Grigoryeva and Ley (2019)

[49]The alternative hypothesis of the test is that the process is a non-linear process, an exponential smooth transition autoregressive (ESTAR) process, which is globally stationary.

test for a ripple effect in house prices in the Vancouver metropolitan area. They find that changes in house prices in both neighbouring and central regions have both a lagged and contemporaneous effect on price changes.

The focus of this thesis is on housing bubbles across London housing markets. There has been some attention in the literature to the ripple affect taking place between London submarkets. For instance Abbott and De Vita (2012) test for convergence of house prices across London boroughs. They do pairwise cointegrating/cotrending tests on the price differentials of all London boroughs. They find that while there is not district level convergence (i.e. of London as a whole), there is "club convergence", where different subdistricts do show convergence. For instance, the City of London shows convergence with all the boroughs that are contiguous with it. The explanation given for this is that the City of London is most exposed to global financial and economic shocks, so will change first, and the surrounding boroughs will be affected through a process of house price contagion.

Bubbles and Coexplosivity

As discussed in the Introduction, and in other chapters of this thesis, a bubble is a period of rapid price increases characterised by self-fulfilling expectations. As the bubble is building, prices need to increase by ever larger amounts in order to maintain the bubble. Because of this, a large part of the literature for testing for bubbles revolves around tests for explosiveness in prices. A very popular testing procedure was developed in Phillips et al. (2011) and Phillips et al. (2015), commonly referred to as PSY, and this methodology is discussed in great detail in earlier chapters of this thesis. Essentially the PSY testing procedure revolves around recursive right tailed unit root tests, and can be used to identify periods where a series was explosive out of a larger sample.

In addition to this, there is an emerging literature on how explosive price series may relate to other series. Nielsen (2010) and Engsted and Nielsen (2012) are the first papers to discuss "coexplosive" processes. In the context Nielsen (2010) and Engsted and Nielsen (2012) use this term, this describes a system with cointegration in a context with explosiveness. The idea is that a process would have several characteristic roots: one is unit root, and one is explosive. The unit root in the system is filtered out by the cointegration, and another filter filters out the explosiveness to get to a system that is stationary. Hence this model can be used in systems where we would expect both cointegration and explosiveness. A key example of this is with rational bubbles - bubbles in a context of rationality. Traditional financial theory and pricing models suggest that asset prices should be cointegrated with their cashflows, and in a bubble prices will be explosive. Hence Engsted and Nielsen (2012) uses the coexplosive model to test for a rational bubble in US stock prices, where stock prices are cointegrated with dividends. This has also been applied to housing markets, with rents replacing dividends, for instance Engsted et al. (2016).

A more recent paper that uses the concept of coexplosiveness is Evripidou et al. (2022). The concept of "coexplosivity", as Evripidou et al. (2022) refer to it, is somewhat simpler than the coexplosiveness used by Nielsen (2010) and Engsted and Nielsen (2012). In this context, coexplosivity is where a linear combination of two explosive processes is stationary, similar to cointegration. Here this is not used to analyse how asset prices relate to fundamentals (as it is unlikely that dividends or rents will be explosive), but is instead used to analyse a system where one bubble will migrate to another market. This migration will likely involve some kind of lag i.e. the bubble will be transmitted after a period of time.

The key idea in this is that the bubble in both markets is essentially the same underlying bubble. The price series are generated by a combination of several different processes, and with coexplosivity a single explosive process drives the bubble in both price series. So a linear combination of these two price series can filter out the bubble. Evripidou et al. (2022) investigate coexplosivity in metals prices, and find evidence of bubble transmission across many different metals. This testing procedure is relatively recent, so it has not been widely implemented. At the time of writing, the only other paper to implement the Evripidou et al. (2022) coexplosivity test is Basse et al. (2023) who find evidence of coexplosivity between the stock market indices the NASDAQ OMX Green Economy Index and the MSCI World Equity Index. So this paper, in which we test for coexplosivity in regional London housing markets, is among the earliest uses of this relatively novel procedure.

iii. Data

Both models use only house prices. The house price data came from the UK House Price Index, which is produced by the Office for National Statistics (ONS). For England, this uses sales data from HM Land Registry. For all the price series, I used log real house prices, deflated by CPI, which is also produced by the ONS. The samples run from January 1998 to December 2011. I focus on this range because in the first paper of this thesis, I find evidence for bubbles in house prices during and before the global financial crisis. In the first paper of this thesis, I found that house price increases post-GFC can be explained by changes in fundamentals (which thus excludes there being a bubble).

iv. Methodology

The Holly et al. (2011) Spatio-Temporal model

In the first part of this paper, I apply the spatio-temporal model of Holly et al. (2011) to London regional housing markets, at the borough level.

This models log real house prices, p_{it}, of $i = 0,1,2,....,N$ regions over time $t = 1,2,....,T$. In this model, there is one region that is treated as the dominant region. Shocks first come to this region, and then propagate out to the other regions. The dominant region is region 0, while the non-dominant regions are regions $i = 1,2,....,N$. Prices for all regions are generated using an error-correction specification[50]. The prices for the dominant region 0 follow a process given by the following equation:

$$\Delta p_{0t} = \phi_{0s}(p_{0,t-1} - \bar{p}^{w}_{0,t-1}) + a_0 + a_{01}\Delta p_{0,t-1} + b_{01}\Delta\bar{p}^{w}_{0,t-1} + \varepsilon_{0t} \tag{86}$$

The rest of the regions follow a process given by the equation:

$$\Delta p_{it} = \phi_{is}(p_{i,t-1} - \bar{p}^{w}_{i,t-1}) + \phi_{i0}(p_{i,t-1} - p_{0,t-1}) + a_i + a_{i1}\Delta p_{i,t-1} + b_{i1}\Delta\bar{p}^{w}_{i,t-1} + c_{i0}\Delta p_{0t} + \varepsilon_{it} \tag{87}$$

For $i = 1,2,....,N$. $\bar{p}^{w}_{it} = \sum_{j=0}^{N} w_{ij}p_{jt}$. The spatial weights, $w_{ij} \geq 0$, are a measure of spatial contiguity with $\sum_{j=1}^{N} w_{ij} = 1$ for $i = 0,1,2,....,N$. Basically these weights measure how connected the other regions are for each region i. Following Holly et al. (2011) I set $w_{ij} = 1/n_i$ if the region j is connected to region i, and zero otherwise (where n_i is the

[50]For simplicity, in this section I use only one lag, but in the practical estimation of the model it is possible to use more than one.

number of regions connected to region i)[51].

The model takes a relatively parsimonious error-correction format. For the dominant region, there is cointegration between it's prices and the surrounding area. This effect is captured by the error-correcting equation $\phi_{0s}(p_{0,t-1} - \bar{p}^w_{0,t-1})$. For each other region, there are two cointegrating relations: between the region and the dominant region, and between the region and the surrounding area. These are captured by the error-correcting equations $\phi_{i0}(p_{i,t-1} - p_{0,t-1})$ and $\phi_{is}(p_{i,t-1} - \bar{p}^w_{i,t-1})$, respectively. A contemporaneous effect of a price change in the dominant region also appears in the price equation for the non-dominant regions, in the $c_{i0}\Delta p_{0t}$ term.

Testing for Endogeneity in Contemporaneous Price Changes in Dominant Region

Note that a $\Delta\bar{p}^w_{0,t}$ term does not appear in the price equation for the dominant region. So there is a contemporaneous effect of price changes in the dominant region on the non-dominant region, but not the other way around. This implies that the contemporaneous price changes in the dominant region are exogenous for the non-dominant regions. This implicit exogeneity can be tested with Wu-Hausman style tests. This involves first estimating an error term from the regression for the dominant region (Equation 86) i.e.

$$\hat{\varepsilon_{0t}} = \Delta p_{0t} - \phi_{0s}(p_{0,t-1} - \bar{p}^w_{0,t-1}) - \hat{a}_0 - \hat{a}_{01}\Delta p_{0,t-1} - \hat{b}_{01}\Delta\bar{p}^w_{0,t-1} \tag{88}$$

And then for each region $i = 1, 2, ..., N$ estimating an auxiliary regression based on the

[51] It could be argued that regions north of the river should not be treated as spatially cotiguous to regions south of the river. I would argue that while such a barrier might have been significant historically, in modern times innovations in communications technology such as the internet make such physical barriers less important.

regression for the non-dominant regions (Equation 87), with the error term $\hat{\varepsilon_{0t}}$ included as an additional explanatory variable i.e.

$$\Delta p_{it} = \phi_{is}(p_{i,t-1} - \bar{p}^{w}_{i,t-1}) + \phi_{i0}(p_{i,t-1} - p_{0,t-1}) + a_i + a_{i1}\Delta p_{i,t-1} + b_{i1}\Delta\bar{p}^{w}_{i,t-1} + c_{i0}\Delta p_{0t} + \lambda_i \hat{\varepsilon_{0t}} + \eta_{it} \tag{89}$$

The intuition behind this is that any endogeneity in Δp_{0t} for Δp_{it}, if there is any, will be captured in $\hat{\varepsilon}_{0t}$, because the terms in Equation 86 are exogenous[52] for Δp_{it}. So if Δp_{0t} were endogenous for Δp_{it}, including $\hat{\varepsilon}_{0t}$ in the regression Equation 87 would affect the estimation of the model and the parameter λ_i will be different from 0 (This would essentially be an IV regression). So a t-test is conducted for the null hypothesis that $\lambda_i = 0$ for each of the regions $i = 1, 2, ..., N$. If λ_i is not significantly different from 0 then our assumption that Δp_{0t} is exogenous for Δp_{it} cannot be rejected.

Spatio-Temporal Impulse response functions We can write the system of equations in a vector/matrix notation.

$$\Delta p_t = a + H p_{t-1} + (A_1 + G_1)\Delta p_{t-1} + C_0 \Delta p_t + \varepsilon_t \tag{90}$$

Where $p_t = (p_{0t}, p_{1t}, ..., p_{Nt})'$, $a = (a_0, a_1, ..., a_N)'$, $\varepsilon_t = (\varepsilon_{0t}, \varepsilon_{1t}, ..., \varepsilon_{Nt})'$ and:

[52]Note that this assumes N>1. Otherwise, if N=1, then $\Delta\bar{p}^{w}_{0,t-1} = \Delta p_{1,t-1}$ and $\Delta\bar{p}^{w}_{1,t-1} = \Delta p_{0,t-1}$, and this procedure cannot be estimated.

$$H = \begin{bmatrix} \phi_{0s} & 0 & \dots & 0 & 0 \\ -\phi_{10} & \phi_{1s}+\phi_{10} & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -\phi_{N-1,0} & 0 & \dots & \phi_{N-1,s}+\phi_{N-1,0} & 0 \\ -\phi_{N0} & 0 & \dots & 0 & \phi_{Ns}+\phi_{N0} \end{bmatrix} - \begin{bmatrix} \phi_{0s}w_0' \\ \phi_{1s}w_1' \\ \vdots \\ \phi_{N-1,s}w_{N-1}' \\ \phi_{Ns}w_N' \end{bmatrix} \tag{91}$$

$$A_1 = \begin{bmatrix} a_{01} & 0 & \dots & 0 & 0 \\ 0 & a_{11} & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & a_{N-1,1} & 0 \\ 0 & 0 & \dots & 0 & a_{N1} \end{bmatrix} \tag{92}$$

$$G_1 = \begin{bmatrix} b_{01}w_0' \\ b_{11}w_1' \\ \vdots \\ b_{N-1,1}w_{N-1}' \\ b_{N1}w_N' \end{bmatrix} \tag{93}$$

$$C_0 = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ c_{10} & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{N-1,0} & 0 & \dots & 0 & 0 \\ c_{N0} & 0 & \dots & 0 & 0 \end{bmatrix} \tag{94}$$

Where $w_i' = (w_{i0}, w_{i1}, ..., w_{iN})$. The matrix H captures the effects of the error correcting terms. The matrix A_1 captures the effect of lagged price changes from a region to itself, while the matrix G_1 captures the effects of lagged price changes in the surrounding regions to each region. The matrix C_0 captures contemporaneous affects from price changes in the dominant region to each other region. Note that this can be solved for contemporaneous price changes - first subtract $C_0 \Delta p_t$ from both sides, and then multiply by $(I_{N+1} - C_0)^{-1}$ to get:

$$\Delta p_t = (I_{N+1} - C_0)^{-1} a + (I_{N+1} - C_0)^{-1} H p_{t-1} + (I_{N+1} - C_0)^{-1} (A_1 + G_1) \Delta p_{t-1} + (I_{N+1} - C_0)^{-1} \varepsilon_t \tag{95}$$

Define $R = (I_{N+1} - C_0)^{-1}$, $\mu = Ra$, $\Pi = RH$ and $\Gamma = R(A_1 + G_1)$. Then we can rewrite Equation 95 as:

$$\Delta p_t = \mu + \Pi p_{t-1} + \Gamma \Delta p_{t-1} + R\varepsilon_t \tag{96}$$

The ranks of a matrix is the number of linearly independent non-zero columns or rows spanned by the matrix. If we define τ_{N+1} to be an N+1 vector of ones, recall that $\sum_{j=1}^{N} w_{ij} = 1$ for $i = 0, 1, 2, ..., N$. This means that $w_i' \tau_{N+1} = 1$ for all $i = 0, 1, 2, ..., N$.

This means that $H\tau_{N+1} = 0$ - each row would just be $\phi_{0s} - \phi_{0s} = 0$, $-\phi_{10} + \phi_{1s} + \phi_{10} - \phi_{1s} = 0$ *et cetera*. Which means that H must be rank deficient. If it were not, and H were full rank, then H would be invertible - but that would mean $H^{-1} H \tau_{N+1} = \tau_{N+1} = H^{-1} 0$. But for any matrix A, $A * 0 = 0$. So H cannot be inverted, which means that H is rank deficient. Given that R is the inverse of $(I_{N+1} - C_0)$, R is invertible and therefore has full rank. It follows that Π is also rank deficient. This means that Equation 96 represents a Vector Error Correction Model (VECM) for p_t, with the number of cointegrating relations being the rank of Π. This means we can rewrite Π as $\Pi = \alpha\beta'$. We can rearrange Equation

96 into a vector autoregression (VAR) format:

$$p_t = \mu + \Phi_1 p_{t-1} + \Phi_2 p_{t-2} + R\varepsilon_t \tag{97}$$

Here $\Phi_1 = I_{N+1} + \alpha\beta' + \Gamma$ and $\Phi_2 = -\Gamma$. We can use this VAR format to conduct impulse response analysis - how each regions house price will respond to a shock in the dominant region's house price. Following Holly et al. (2011) I use Pesaran and Shin (1998) Generalised Impulse Response Functions (GIRFs). The equation for the GIRFs has the following form:

$$g_i(h) = \frac{\Psi_h R \Sigma e_i}{\sqrt{\sigma_{ii}}} \tag{98}$$

Where

$$\Psi_h = \Phi_1 \Psi_{h-1} + \Phi_2 \Psi_{h-2} \text{for } h = 0, 1, ... \tag{99}$$

with initial values $\Psi_0 = I_{N+1}$ and $\Psi_h = 0$ for $h < 0$. e_i is an $(N+1) \times 1$ vector of zeroes, with the exception of the ith element which is one.

$$\Sigma = \begin{bmatrix} \sigma_{01} & 0 & 0 & \dots & 0 & 0 \\ 0 & \sigma_{11} & \sigma_{12} & \dots & \sigma_{1,N-1} & \sigma_{1N} \\ 0 & \sigma_{21} & \sigma_{22} & \dots & \sigma_{2,N-1} & \sigma_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \sigma_{N-1,1} & \sigma_{N-1,2} & \dots & \sigma_{N-1,N-1} & \sigma_{N-1,N} \\ 0 & \sigma_{N1} & \sigma_{N2} & \dots & \sigma_{N,N-1} & \sigma_{NN} \end{bmatrix} \tag{100}$$

Where $\sigma_{ij} = E(\varepsilon_{it}\varepsilon_{jt})$. Each of these can be estimated consistently via OLS residuals i.e. $\hat{\sigma}_{00} = T^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_{0t}^{2}$ and $\hat{\sigma}_{ij} = T^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_{it}\hat{\varepsilon}_{jt}$ for $i, j = 1, 2, ..., N$.

Geographical Regions The London boroughs are shown in Figure 40. The boroughs are as follows: 1. City of London 2. City of Westminster 3. Kensington & Chelsea 4. Hammersmith & Fulham 5. Wandsworth 6. Lambeth 7. Southwark 8. Tower Hamlets 9. Hackney 10. Islington 11. Camden 12. Brent 13. Ealing 14. Hounslow 15. Richmond upon Thames 16. Kingston upon Thames 17. Merton 18. Sutton 19. Croydon 20. Bromley 21. Lewisham 22. Greenwich 23. Bexley 24. Havering 25. Barking and Dagenham 26. Redbridge 27. Newham 28. Waltham Forest 29. Haringey 30. Enfield 31. Barnet 32. Harrow 33. Hillingdon. In the spatio-temporal model, one region is assumed to be the dominant region: shocks to prices hit this region first, and then radiate out from this dominant region to the other regions across space and time. For the purposes of this model, I aggregated together the most central London boroughs into one region "Central London", aggregated based on transaction data used in the HPI published by the ONS. The London boroughs which went into the Central London region are highlighted in cyan in Figure 40. These were The City of London, Camden, Islington, Hackney, Tower Hamlets, Southwark, Lambeth, and The City of Westminster,

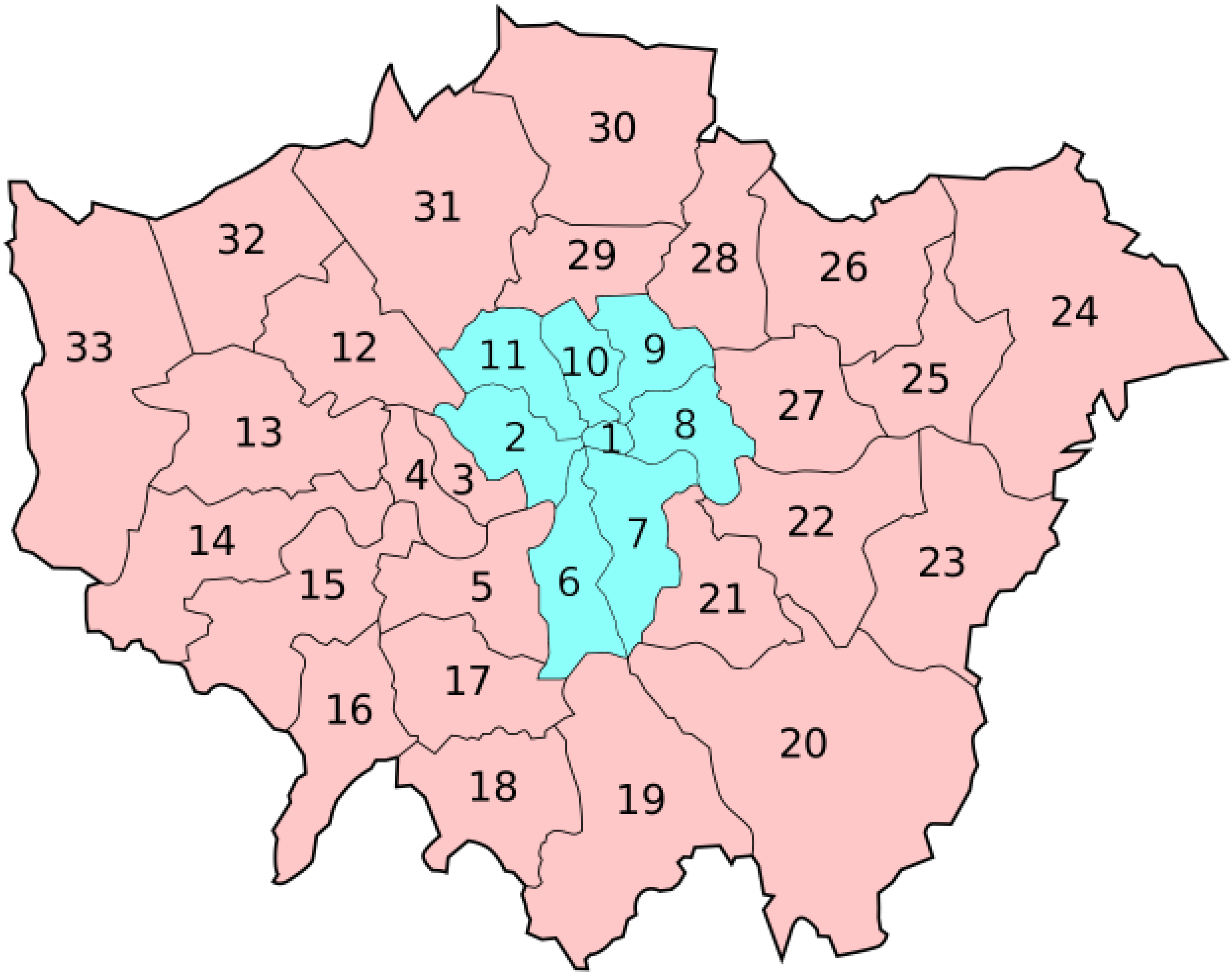


Figure 40: *London boroughs, with Central London boroughs indicated in cyan*

The Evripidou et al. (2022) Coexplosive Model

Evripidou et al. (2022) analyse a system where there are two explosive processes that are fundamentally linked. They call this "co-explosivity". In the case of price bubbles, this would be where the bubble migrates to different assets.

The Explosive Data Generation Process Consider the following data generation process:

$$x_t = \mu_x + u_{x,t} \tag{101}$$

Where

$$u_{x,t} = \begin{cases} u_{x,t-1} + \varepsilon_{x,t} & t = 2, ..., [\tau_{x,1}T] \\ (1+\delta_{x,1})u_{x,t-1} + \varepsilon_{x,t} & t = [\tau_{x,1}T]+1, ..., [\tau_{x,2}T] \\ (1-\delta_{x,2})u_{x,t-1} + \varepsilon_{x,t} & t = [\tau_{x,2}T]+1, ..., [\tau_{x,3}T] \\ u_{x,t-1} + \varepsilon_{x,t} & t = [\tau_{x,3}T]+1, ..., T \end{cases} \tag{102}$$

Here μ_x is a constant, $\tau_{x,1}$, $\tau_{x,2}$ and $\tau_{x,3}$ are fractions of the overall sample, and $\delta_{x,1}$ and $\delta_{x,2}$ are constants >0. So basically the variable x_t will be a unit root process, up until time $\tau_{x,1}T$. At this point the process becomes explosive. This ends at time $\tau_{x,2}T$ where the process becomes stationary. This ends at time $\tau_{x,3}T$, after which point the process becomes unit root again. This is a general data generation process for describing a bubble period. Supposing x_t were a price series, at first it is a unit root process, i.e. a random walk, as most price series tend to be. Then the bubble starts - the price starts to increase at an accelerating rate - which is captured by when the series becomes explosive. The bubble collapses - this is where the series is stationary. After which the bubble episode is over, and the price series goes back to a normal random walk. Note that this process can be generalised by varying $\tau_{x,1}$, $\tau_{x,2}$ and $\tau_{x,3}$ to specify different types of regimes. For instance, if $\tau_{x,1} = \tau_{x,2} = \tau_{x,3}$ then this is no bubble period, if $\tau_{x,1} < \tau_{x,2} = \tau_{x,3} < 1$ then there is no period when the bubble collapses - when the bubble ends, the process goes back to being a random walk.

The Coexplosive Model Suppose there are two processes, y_t and x_t. x_t was generated by a process such as Equation 102. y_t is generated by the following process:

$$y_t = \mu_y + \beta_x x_{t-i} + \beta_z z_t + \varepsilon_{y,t} \tag{103}$$

Here z_t is another process generated by Equation 102 (with all the x subscripts replaced by z subscripts). $\varepsilon_{y,t}$ is an I(0) process, $E(\varepsilon_{y,t}) = 0$. If $\beta_x = 0$ and $\beta_z > 0$, then both y_t and x_t are explosive, and the two processes are unrelated to each other. If $\beta_x > 0$ and $\beta_z = 0$, then both processes are explosive, and they are fundamentally related to each other. In this case, a linear combination of y_t and x_t, $y_t - \mu_y - \beta_x x_{t-i} = \varepsilon_{y,t}$, is I(0). Similar to cointegration, following Evripidou et al. (2022) we can call y_t and x_t "coexplosive" in this case. They are two explosive processes, but the explosive part of y_t ultimately comes from another explosive process, x_t. Note that there is a possibility of a lag, i, between y_t and x_t.

We can test for this coexplosivity between y_t and x_t by testing for stationarity of some linear combination of y_t and x_{t-i}. We can do this with a KPSS style statistic i.e.

$$S = \hat{\sigma}_y^{-2}(T - |i|)^{-2} \sum_{t=i1(i>0)+1}^{T+i1(i<0)} \left(\sum_{s=i1(i>0)+1}^{t} \hat{e}_{y,s} \right)^2 \tag{104}$$

Where $\hat{e}_{y,t}$ is the residual from a regression of y_t on x_{t-i} i.e. $\hat{e}_{y,s} = y_t - \hat{\mu}_y - \hat{\beta}_x x_{t-i}$. The KPSS statistic is thus derived from the sum of the cumulative sums of residuals. The intuition behind this is that if these residuals are stationary, then the cumulative sum of the residuals will not deviate significantly from zero. If the process is unit root, then the cumulative sum of residuals might differ significantly from zero at some point (before it necessarily returns to zero by the end of the sample).

Note that we make no assumptions regarding heteroscedasticity or the covariance of model errors. Critical values are obtained via a wild bootstrap procedure, following Evripidou et al. (2022).

If there exists a linear combination of the two explosive series that is stationary, this is interpreted as evidence of coexplosivity - the explosive series are fundamentally related to each other. If the explosive series are caused by an asset price bubble, this could be interpreted as evidence of bubble contagion - that one bubble has been transmitted from one asset class to another.

Lag Selection In practise the lag, i, is not known a priori. In practice, the coexplosivity regression is estimated for a range of lags, chosen a priori to be a reasonable maximum range. For each lag, j, we calculate $\hat{\sigma}_{y,j}^{-2} = (T-|i|)^{-1}\sum_{t=i1(i>0)+1}^{T+i1(i<0)} \hat{e}_{y,t,j}^{2}$. The lag which we use for the KPSS test is then the lag which has the smallest value of $\hat{\sigma}_{y,j}^{-2}$. More formally, i is estimated by $\hat{i} = \arg\min_{j \in J} \hat{\sigma}_{y,j}^{-2}$, where J is the set of values used. The intuition behind this is that if the lag term is chosen incorrectly, then $\hat{e}_{y,t,j}$ will have an explosive component. Hence the magnitude of the $\hat{e}_{y,t,j}$ terms will be larger. Hence we can estimate the correct lag by minimising $\hat{\sigma}_{y,j}^{-2}$.

The Wild Bootstrap Procedure The wild bootstrap procedure used to construct critical values proceeds as follows:

Step 1: With the residuals (for which we are testing stationarity), $\hat{e}_{y,t}$, construct a bootstrap sample: $y_{tb} = w_t\hat{e}_{y,t}$ for $t = i1(i<0)+1, ..., T+i1(i<0)$, where w_t is *IID* and N(0,1).

Step 2: y_{tb} is regressed on a constant and x_{t-i}. The residuals from this regression are denoted

$\hat{e}_{tb}$, and the KPSS statistic, S_b is calculated as:

$$S = \hat{\sigma}_b^{-2}(T-|i|)^{-2}\sum_{t=i1(i>0)+1}^{T+i1(i<0)}\left(\sum_{s=i1(i>0)+1}^{t}\hat{e}_{sb}\right)^2 \tag{105}$$

Where $\hat{\sigma}_b^{-2} = (T-|i|)^{-1}\sum_{t=i1(i>0)+1}^{T+i1(i<0)}\hat{e}_{tb}^2$

Step 3: Steps 1 & 2 are repeated for $b = 1, 2, ..., B$ bootstrapped series.

Step 4: The α significance level critical value is then the $1-\alpha$ quantile of the empirical distribution of S_b.

The PSY Procedure One issue with the Evripidou et al. (2022) procedure is that we have to ensure the time series we are using are both explosive. Otherwise the linear combination of the two being stationary is indicative of cointegration, not coexplosivity. To find time periods that are explosive, I use Phillips et al. (2011) and Phillips et al. (2015) (PSY) recursive right tailed unit root tests, as in the other two papers of this thesis. This procedure involves repeatedly calculating Augmented Dickey-Fuller statistics (ADF) over recursive timeframes. Consider a standard ADF regression, shown in Equation 106. If the process y_t is explosive, this can be tested with a t-statistic on the parameter δ (this is the ADF statistic). If $\delta > 0$, then the process is explosive.

$$\Delta y_t = \delta y_{t-1} + \sum_{i=1}^{k}\alpha_i \Delta y_{t-i} + \varepsilon_t \tag{106}$$

The PSY procedure then involves calculating ADF statistics repeatedly, over different time periods. We assume that there is some minimum sample length, denoted r_0. Then for all start dates and end dates, an ADF statistic is calculated. The GSADF statistic is then the supremum of these statistics:

$$GSADF_{r_2}(r_0) = \sup_{r_1 \in [1, r_2 - r_0], r_2 \in [r_0, T]} ADF_{r_1}^{r_2} \tag{107}$$

v. Results

The Spatio-Temporal Model

Results for the estimation of the Spatio-Temporal model are given in Tables 22, 23 and 24. Clearly the results for the first error correction term, which captures convergence between the region and the dominant region, was not statistically significant for almost all the regions, apart from Hammersmith & Fulham and Wandsworth. This means that there isn't much evidence for price convergence between Central London and the other London boroughs. Likewise, generally speaking a lagged price change in Central London only has a significant effect in three regions: Brent, Croydon, and Kingston upon Thames [53].

The second error correction term, which captures convergence between the region and the average of adjacent regions, had far more significant results. For many of the regions, the term is statistically significant. This is for the regions Barking & Dagenham, Ealing, Greenwich, Haringey, Hillingdon, Kingston upon Thames, Merton, Richmond upon Thames, Sutton, and Waltham Forest. Note how the term is always negative. The coefficient being negative means that when this region's prices are higher than its neighbours (the price difference is positive), the prices in this region will most likely decrease next period. And vice versa: If the regions prices are lower than its neighbours (the price difference is negative), then the next periods price change is likely to be positive. This is evidence of

[53] The first error correction term and the lagged effect of price changes in Central London are only reported where statistically significant

Region	EC1	t-ratio	EC2	t-ratio	Own Lag Effects	t-ratio
Central London	-	-	-	-	0.233	1.796
Barking and Dagenham	-	-	-0.035	-1.889	0.552	4.967
Barnet	-	-	-	-	-0.203	-1.365
Bexley	-	-	-	-	0.285	2.502
Brent	-	-	-	-	0.231	2.221
Bromley	-	-	-	-	0.041	0.363
Croydon	-	-	-	-	0.614	6.574
Ealing	-	-	-0.249	-4.209	0.45	4.338
Enfield	-	-	-	-	0.486	4.753
Greenwich	-	-	-0.158	-3.865	0.324	3.976
Hammersmith	-0.128	-3.291	-	-	0.095	0.698
Haringey	-	-	-0.12	-3.281	0.119	1.49
Harrow	-	-	-	-	-0.383	-2.867
Havering	-	-	-	-	0.284	3.572
Hillingdon	-	-	-0.078	-2.477	0.555	5.653
Hounslow	-	-	-	-	-0.096	-0.724
Kensington and Chelsea	-	-	-	-	-0.244	-2.158
Kingston upon Thames	-	-	-0.129	-3.347	0.014	0.112
Lewisham	-	-	-	-	-0.003	-0.03
Merton	-	-	-0.324	-5.247	0.121	1.074
Newham	-	-	-	-	-0.139	-0.836
Redbridge	-	-	-	-	0.349	2.962
Richmond upon Thames	-	-	-0.094	-2.911	0.423	3.814
Sutton	-	-	-0.106	-2.535	0.563	4.718
Waltham Forest	-	-	-0.042	-1.917	0.355	3.45
Wandsworth	-0.359	-6.076	-	-	0.621	5.86

Table 22: *Spatio-Temporal model results (i)*

local convergence between regional prices in London boroughs.

In addition to the second error correction term, the lagged effect of neighbouring prices changes is also significant for many regions: Central London, Barnet, Bexley, Bromley, Greenwich, Hammersmith, Harrow, Havering, Hounslow, Lewisham, Newham, Redbridge, and Waltham Forest. For all of the regions where this effect is statistically significant, the effect is positive: so a price increase in neighbouring regions will increase the price in the region in the next period. Clearly then, in addition to spatial convergence, there is also more general spatial dependence across London boroughs.

I would argue that the results show clear evidence of a ripple effect across London boroughs. Boroughs show both regional convergence and spatial dependence: if there is a price differential between neighbouring regions, there is a tendency for the prices to equalise: at the same time, if there are price changes in neighbouring regions then the region's prices will also tend to change in the same direction next period. There is not evidence for Central London acting as a dominant region however. There is not evidence for shocks hitting Central London first, and then radiating out to the rest of London. Figure 41 visualises the Generalised Impulse Response Functions of how house prices in the boroughs change over time following a one unit shock to Central London house prices. The boroughs are ordered in distance from Central London (in terms of no of boroughs away): however there doesn't seem to be much difference between the regions that are closer to Central London as opposed to those farther away.

Region	Neighbour Lag Effects	t-ratio	Central London Lag Effects	t-ratio	Central London Contemp Effects	t-ratio
Central London	0.521	4.356	-	-	-	-
Barking and Dagenham	0.171	1.228	-	-	0.211	2.837
Barnet	0.73	4.66	-	-	0.35	5.012
Bexley	0.238	2.049	-	-	0.342	5.496
Brent	-0.119	-1.187	0.355	3.607	0.396	5.298
Bromley	0.532	4.624	-	-	0.329	6.658
Croydon	-0.142	-1.044	0.211	2.393	0.294	5.278
Ealing	0.128	1.292	-	-	0.503	7.366
Enfield	0.087	1.08	-	-	0.258	3.932
Greenwich	0.231	2.141	-	-	0.313	3.928
Hammersmith	0.375	2.751	-	-	0.667	5.649
Haringey	-0.026	-0.19	-	-	0.859	8.505
Harrow	0.917	6.241	-	-	0.372	6.111
Havering	0.306	3.7	-	-	0.236	3.981
Hillingdon	0.028	0.334	-	-	0.221	3.629
Hounslow	0.36	2.898	-	-	0.511	6.062
Kensington and Chelsea	0.228	0.988	-	-	0.662	3.182
Kingston upon Thames	-0.09	-0.759	0.702	4.444	0.431	5.377
Lewisham	0.526	4.339	-	-	0.355	4.477
Merton	0.207	1.562	-	-	0.441	5.482
Newham	0.919	4.65	-	-	0.324	3.77
Redbridge	0.224	2.189	-	-	0.295	4.242
Richmond upon Thames	0.112	1.053	-	-	0.446	4.312
Sutton	0.034	0.328	-	-	0.348	5.153
Waltham Forest	0.244	2.133	-	-	0.33	4.371
Wandsworth	-0.018	-0.168	-	-	0.795	7.519

Table 23: *Spatio-Temporal model results (ii)*

Region	Wu-Haus	R^2	k_own	k_neib	k_Base
Central London	-	0.411	4	1	-
Barking and Dagenham	-0.517	0.527	4	1	0
Barnet	-0.622	0.59	4	3	0
Bexley	0.516	0.456	2	1	0
Brent	-1.682	0.497	4	1	1
Bromley	-0.664	0.613	3	3	0
Croydon	1.181	0.609	4	1	1
Ealing	0.045	0.558	2	1	0
Enfield	-1.26	0.488	4	1	0
Greenwich	1.565	0.399	1	1	0
Hammersmith	-0.46	0.48	4	1	0
Haringey	0.24	0.451	1	1	0
Harrow	-0.466	0.581	3	3	0
Havering	-0.482	0.469	1	1	0
Hillingdon	0.953	0.514	4	1	0
Hounslow	-1.584	0.535	4	1	0
Kensington and Chelsea	0.582	0.293	3	1	0
Kingston upon Thames	-2.188	0.599	3	1	4
Lewisham	-1.285	0.454	3	1	0
Merton	-0.673	0.565	3	1	0
Newham	1.182	0.503	4	3	0
Redbridge	0.041	0.509	4	1	0
Richmond upon Thames	-1.105	0.454	4	1	0
Sutton	-0.392	0.521	4	1	0
Waltham Forest	0.611	0.546	4	1	0
Wandsworth	1.205	0.463	2	1	0

Table 24: *Spatio-Temporal model results (iii)*

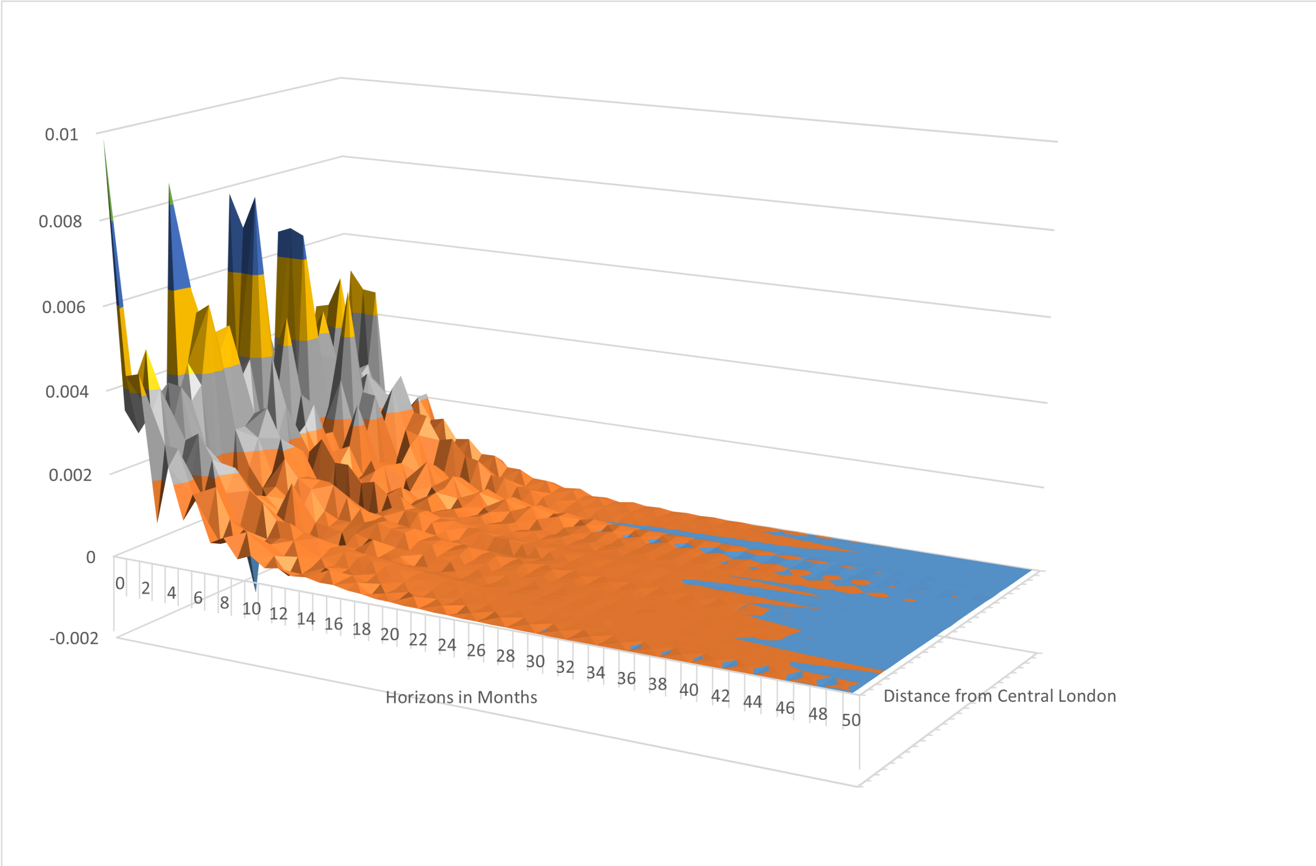


Figure 41: *Generalised impulse responses of a one unit shock to Central London house prices, showing house price changes across time and across regions*

Now while there are little lagged effects from Central London prices, the contemporaneous effects are much more significant, with every single borough having a statistically significant effect from contemporaneous price changes in Central London. Note that almost none of the Wu-Hausman statistics are statistically significant at 5%, so simultaneity bias won't have affected the estimation much (Kingston upon Thames is the only borough with a Wu-Hausman statistic above 1.96 in magnitude however, so perhaps the results for this borough can be dismissed due to simultaneity bias). The parameter for the effect of a contemporaneous price increase in Central London is positive for every region - so an increase in Central London prices will contemporaneously increase prices in other regions.

One explanation for the lack of temporal lag between Central London Price might be that as London is relatively smaller than the rest of the UK, shocks transmit so quickly that it is contemporaneous (when data is monthly). It might be that price shocks do hit Central London first and then radiate out, it is just that the data is not frequent enough to capture this, so it just shows as a contemporaneous effect. Unfortunately it would be difficult to see this from the data, as it is rare to get home price data that is more frequent than monthly.

Explosive Sample Timeframes

I report the GSADF statistics and corresponding 5% critical values, along with the time-frames of peak explosiveness in Table 25. Clearly for all regions, there was a period of explosive prices. It is important that the entire sample is explosive for both series in which the Evripidou et al. (2022) test procedure is done. From Table 25, there are four sets of explosive periods that almost all regions fall into (excluding Ealing). The first includes February 1998 until January 2000. This is for the regions: Bexley, Bromley, Croydon, Hounslow, Newham, Redbridge, Sutton, and Waltham Forest. Note that for some of these regions, the period of explosive starts earlier than February 1998 or ends later than January 2000. I do this time-frame so that for all of these regions, the sample is a subsample of an explosive period. The second period of explosive behaviour runs from February 2000 until July 2002. This is for the regions Barking and Dagenham, Barnet, and Harrow. The third period of explosive behaviour runs from January 2005 until July 2007. This is for the regions Central London, Brent, Enfield, Hammersmith, Haringey, Kensington and Chelsea, Merton, Richmond upon Thames, and Wandsworth. The fourth period of explosive behaviour runs from February 2007 until December 2008. This is for the regions Greenwich, Havering, Hillingdon, Kingston upon Thames, and Lewisham. So I run the Evripidou et al. (2022) procedure four times, for the four periods of explosive price behaviour.

The Coexplosive Model

The results of the coexplosive model compliment the results of the spatio-temporal model. The results are given in Table 26 and Table 27 for the regions with the explosive period running from February 1998 until January 2000, Table 28 for the regions with the explosive period running from February 2000 until July 2002, Table 29, Table 30, Table 31, and Table 32 for the regions with the explosive period running from January 2005 until July 2007, and Table 33 for the regions with the explosive period running from February 2007 until December 2008. For two of these periods, there was some coexplosivity between regions. For the period running from 1998 to 2000, the regional pairs which had coexplosivity, significant at the 5% level, were: Newham and Bexley, Croydon and Waltham Forest, Newham and Redbridge, Newham and Sutton, Waltham Forest and Redbridge. Note that the coexplosivity between Newham and Bexley, and Croydon and Waltham Forest were not contemporaneous. A linear combination of one region's prices, and the other region's lagged prices, was stationary.

Some of this coexplosivity took place in contiguous regions i.e. Newham & Redbridge, Redbridge & Waltham Forest. Likewise Bexley is broadly in the same part of London as Newham. For all of these pairs, one or both of the regions had neighbour lag effects or convergence effects in the spatio-temporal model: Waltham had evidence of regional convergence, and Newham, Redbridge and Waltham Forest all have significant neighbour

Region	Start Date	End Date	GSADF test statistic	5% Critical Value
Central London	01-Feb-04	01-Jul-07	5.610755345	1.576828934
Barking and Dagenham	01-Feb-00	01-Jan-03	5.002171072	1.687059383
Barnet	01-Feb-00	01-Jul-02	3.094022371	1.680562523
Bexley	01-Jan-98	01-Aug-00	3.731824173	1.689625816
Brent	01-Mar-04	01-Jul-07	3.25298939	1.589059676
Bromley	01-Feb-98	01-Apr-00	3.676879164	1.795424142
Croydon	01-Jan-98	01-Jan-00	4.205314791	1.791286321
Ealing	01-Aug-02	01-Aug-04	2.896416968	1.549113356
Enfield	01-Jan-04	01-Aug-07	4.380252936	1.799587952
Greenwich	01-Nov-06	01-May-09	2.176058563	1.65381131
Hammersmith	01-Jan-04	01-Sep-07	1.911602372	1.454051292
Haringey	01-Mar-04	01-Sep-07	2.547480056	1.722611563
Harrow	01-Jan-00	01-Sep-02	4.238144081	1.936144407
Havering	01-Dec-06	01-Dec-08	4.303138875	1.70943825
Hillingdon	01-Jan-07	01-Feb-09	3.616713801	1.658387981
Hounslow	01-Jan-98	01-Jan-00	2.010982778	1.701805621
Kensington and Chelsea	01-Jan-05	01-Sep-07	2.195367345	1.591109972
Kingston upon Thames	01-Feb-07	01-Mar-09	5.455307008	1.498148466
Lewisham	01-Dec-06	01-Dec-08	3.19088559	1.727777153
Merton	01-Jul-04	01-Aug-07	3.840133092	1.723792371
Newham	01-Jan-98	01-Jul-00	3.667686188	1.809850129
Redbridge	01-Jan-98	01-Jan-00	5.85890336	1.689148269
Richmond upon Thames	01-Dec-04	01-Aug-07	3.907768716	1.81082178
Sutton	01-Jan-98	01-Jan-00	4.713629166	1.781041472
Waltham Forest	01-Jan-98	01-Jan-00	3.221256769	1.62806001
Wandsworth	01-Feb-04	01-Sep-07	2.762439758	1.576071006

Table 25: *PSY results*

lag effects.

With the period of explosiveness running from 2005 until 2007, there are several regions that show signs of coexplosivity at the 5% significance level: Central London and Brent, Kensington & Chelsea and Central London, Merton and Central London, Brent and Enfield, Brent and Wandsworth, Hammersmith & Fulham and Haringey. All of these were non-contemporaneous relations.

Much of this coexplosivity took place in contiguous regions: Central London and Brent, Kensington & Chelsea and Central London, Merton and Central London, Hammersmith & Fulham and Haringey are all contiguous. Likewise Brent & Enfield and Brent & Wandsworth are in similar areas of London. Again, for all of these pairs, one or both of the regions had neighbour lag effects or convergence effects in the spatio-temporal model: Brent, Hammersmith & Fulham, Merton all have evidence of regional convergence, and Central London have Hammersmith & Fulham neighbour lag effects.

What that means is that prices in these regions went through periods of explosiveness, and these periods of explosive house price behaviour were fundamentally related to a separate period of explosive house price behaviour that had already happened in another region. Bubbles transmitted from one borough to another, and in most cases this occurred in geographically proximate boroughs, in the same boroughs which had evidence of convergence and ripple effects in the spatio-temporal model.

vi. Concluding Remarks

In this paper I have found strong evidence of local price convergence and spatial dependence at the borough level in housing markets across Greater London. In other words, there is

y	x	Lag value[a]	Test Statistic	2.5% Critical Value	5% Critical Value	10% Critical Value	P-value
Bexley	Croydon	0	0.145	0.215	0.187	0.151	0.112
Croydon	Bexley	0	0.189	0.278	0.252	0.203	0.12
Bexley	Newham	1	0.154	0.212	0.189	0.157	0.108
Newham	Bexley	-1	0.202	0.238	0.2	0.161	**0.05**
Bexley	Redbridge	1	0.146	0.269	0.223	0.185	0.152
Redbridge	Bexley	-1	0.145	0.246	0.216	0.169	0.16
Bexley	Waltham Forest	0	0.118	0.181	0.158	0.138	0.174
Waltham Forest	Bexley	-1	0.175	0.212	0.189	0.16	0.072
Bromley	Hounslow	0	0.191	0.273	0.226	0.186	0.094
Hounslow	Bromley	-1	0.161	0.192	0.174	0.149	0.074
Bromley	Newham	0	0.212	0.285	0.24	0.196	0.082
Newham	Bromley	0	0.172	0.22	0.189	0.165	0.082
Croydon	Newham	0	0.17	0.252	0.213	0.174	0.102
Newham	Croydon	-1	0.173	0.256	0.234	0.185	0.106
Croydon	Redbridge	1	0.225	0.263	0.24	0.205	0.064
Redbridge	Croydon	-2	0.209	0.243	0.214	0.185	0.052
Croydon	Sutton	-1	0.18	0.322	0.271	0.215	0.156
Sutton	Croydon	0	0.176	0.25	0.224	0.191	0.12
Croydon	Waltham Forest	1	0.222	0.24	0.193	0.161	**0.032**
Waltham Forest	Croydon	-1	0.159	0.224	0.187	0.158	0.098

Table 26: *Coexplosive model, with explosive period 1998 - 2000*

[a] Here, and with subsequent tables, a negative value refers to lag, and a positive value refers to a lead.

y	x	Lag value	Test Statistic	2.5% Critical Value	5% Critical Value	10% Critical Value	P-value
Newham	Redbridge	0	0.239	0.258	0.225	0.192	**0.034**
Redbridge	Newham	0	0.19	0.235	0.2	0.177	0.076
Newham	Sutton	0	0.244	0.254	0.231	0.188	**0.036**
Sutton	Newham	0	0.195	0.228	0.205	0.168	0.064
Newham	Waltham Forest	0	0.144	0.215	0.184	0.156	0.12
Waltham Forest	Newham	0	0.178	0.23	0.196	0.159	0.074
Redbridge	Waltham Forest	0	0.19	0.22	0.196	0.169	0.064
Waltham Forest	Redbridge	0	0.232	0.258	0.229	0.195	**0.048**

Table 27: *Coexplosive model, with explosive period 1998 - 2000, continued*

y	x	Lag value	Test Statistic	2.5% Critical Value	5% Critical Value	10% Critical Value	P-value
Barking and Dagenham	Barnet	0	0.182	0.232	0.195	0.159	0.06
Barnet	Barking and Dagenham	0	0.143	0.219	0.178	0.145	0.114
Barking and Dagenham	Harrow	1	0.157	0.206	0.18	0.148	0.084
Harrow	Barking and Dagenham	-1	0.168	0.215	0.185	0.153	0.068
Barnet	Harrow	0	0.126	0.204	0.18	0.146	0.156
Harrow	Barnet	0	0.162	0.242	0.202	0.16	0.096

Table 28: *Coexplosive model, with explosive period 2000 - 2002*

y	x	Lag value	Test Statistic	2.5% Critical Value	5% Critical Value	10% Critical Value	P-value
Central London	Brent	-10	0.173	0.211	0.172	0.143	**0.05**
Brent	Central London	-7	0.129	0.157	0.136	0.119	0.064
Central London	Enfield	-10	0.155	0.182	0.165	0.142	0.072
Enfield	Central London	4	0.2	0.267	0.231	0.2	0.098
Central London	Kensington and Chelsea	-6	0.143	0.241	0.203	0.165	0.146
Kensington and Chelsea	Central London	-10	0.219	0.22	0.197	0.172	**0.032**
Central London	Merton	-6	0.15	0.226	0.2	0.153	0.106
Merton	Central London	3	0.169	0.192	0.165	0.134	**0.046**
Central London	Richmond upon Thames	-6	0.135	0.209	0.182	0.148	0.142
Richmond upon Thames	Central London	-2	0.124	0.218	0.191	0.157	0.18
Central London	Wandsworth	0	0.12	0.258	0.214	0.181	0.248
Wandsworth	Central London	-1	0.094	0.239	0.192	0.159	0.366
Brent	Enfield	-10	0.192	0.208	0.178	0.146	**0.034**
Enfield	Brent	-6	0.2	0.281	0.242	0.193	0.086
Brent	Haringey	-7	0.137	0.165	0.146	0.122	0.062
Haringey	Brent	-10	0.175	0.197	0.177	0.147	0.058
Brent	Kensington and Chelsea	-6	0.129	0.155	0.139	0.119	0.074
Kensington and Chelsea	Brent	-10	0.244	0.371	0.328	0.273	0.132
Brent	Merton	-8	0.132	0.16	0.136	0.115	0.056
Merton	Brent	-6	0.184	0.292	0.23	0.182	0.098

Table 29: *Coexplosive model, with explosive period 2005 - 2007*

y	x	Lag value	Test Statistic	2.5% Critical Value	5% Critical Value	10% Critical Value	P-value
Brent	Richmond upon Thames	-6	0.114	0.18	0.155	0.13	0.146
Richmond upon Thames	Brent	-6	0.274	0.33	0.29	0.241	0.068
Brent	Wandsworth	-8	0.151	0.167	0.138	0.11	**0.038**
Wandsworth	Brent	-5	0.138	0.249	0.192	0.164	0.184
Enfield	Haringey	-10	0.163	0.22	0.205	0.17	0.128
Haringey	Enfield	-5	0.172	0.249	0.222	0.175	0.104
Enfield	Kensington and Chelsea	-10	0.159	0.186	0.162	0.137	0.058
Kensington and Chelsea	Enfield	-7	0.23	0.382	0.328	0.263	0.128
Enfield	Merton	-2	0.183	0.337	0.255	0.195	0.118
Merton	Enfield	-10	0.161	0.26	0.224	0.189	0.148
Enfield	Richmond upon Thames	-2	0.162	0.261	0.237	0.189	0.14
Richmond upon Thames	Enfield	-6	0.179	0.314	0.278	0.215	0.172
Enfield	Wandsworth	-10	0.158	0.21	0.178	0.149	0.08
Wandsworth	Enfield	-5	0.155	0.351	0.289	0.232	0.248
Hammersmith and Fulham	Haringey	-10	0.17	0.172	0.154	0.136	**0.026**
Haringey	Hammersmith and Fulham	-1	0.15	0.238	0.208	0.174	0.154
Hammersmith and Fulham	Kensington and Chelsea	-8	0.137	0.177	0.161	0.142	0.118
Kensington and Chelsea	Hammersmith and Fulham	-10	0.271	0.326	0.285	0.24	0.064
Hammersmith and Fulham	Merton	-10	0.149	0.181	0.163	0.14	0.074
Merton	Hammersmith and Fulham	-10	0.161	0.26	0.224	0.191	0.166

Table 30: *Coexplosive model, with explosive period 2005 - 2007, continued*

y	x	Lag value	Test Statistic	2.5% Critical Value	5% Critical Value	10% Critical Value	P-value
Hammersmith and Fulham	Richmond upon Thames	-9	0.143	0.169	0.151	0.128	0.064
Richmond upon Thames	Hammersmith and Fulham	-10	0.198	0.229	0.202	0.17	0.056
Hammersmith and Fulham	Wandsworth	-10	0.144	0.16	0.148	0.127	0.058
Wandsworth	Hammersmith and Fulham	-1	0.133	0.278	0.238	0.195	0.27
Haringey	Merton	-6	0.133	0.146	0.135	0.114	0.052
Merton	Haringey	-10	0.155	0.251	0.222	0.192	0.166
Haringey	Richmond upon Thames	-6	0.133	0.175	0.147	0.122	0.068
Richmond upon Thames	Haringey	-8	0.14	0.18	0.153	0.136	0.09
Haringey	Wandsworth	0	0.13	0.266	0.235	0.181	0.242
Wandsworth	Haringey	0	0.225	0.291	0.236	0.199	0.066
Kensington and Chelsea	Merton	-10	0.199	0.262	0.224	0.198	0.096
Merton	Kensington and Chelsea	-10	0.157	0.235	0.211	0.174	0.122

Table 31: *Coexplosive model, with explosive period 2005 - 2007, continued*

y	x	Lag value	Test Statistic	2.5% Critical Value	5% Critical Value	10% Critical Value	P-value
Kensington and Chelsea	Richmond upon Thames	-10	0.19	0.253	0.218	0.18	0.086
Richmond upon Thames	Kensington and Chelsea	-10	0.156	0.171	0.153	0.127	0.042
Merton	Richmond upon Thames	3	0.148	0.288	0.231	0.174	0.148
Richmond upon Thames	Merton	-3	0.22	0.397	0.306	0.231	0.124
Merton	Wandsworth	-10	0.151	0.236	0.208	0.177	0.142
Wandsworth	Merton	-3	0.204	0.244	0.208	0.182	0.058
Richmond upon Thames	Wandsworth	-1	0.099	0.255	0.213	0.169	0.316
Wandsworth	Richmond upon Thames	0	0.108	0.256	0.207	0.164	0.302

Table 32: *Coexplosive model, with explosive period 2005 - 2007, continued*

y	x	Lag value	Test Statistic	2.5% Critical Value	5% Critical Value	10% Critical Value	P-value
Greenwich	Havering	1	0.277	0.352	0.289	0.23	0.06
Havering	Greenwich	-1	0.218	0.355	0.278	0.228	0.116
Greenwich	Kingston upon Thames	-1	0.156	0.444	0.369	0.291	0.316
Kingston upon Thames	Greenwich	1	0.193	0.483	0.413	0.335	0.288
Havering	Kingston upon Thames	-1	0.289	0.354	0.295	0.236	0.052
Kingston upon Thames	Havering	1	0.215	0.394	0.302	0.235	0.126
Havering	Lewisham	-1	0.26	0.443	0.36	0.267	0.108
Lewisham	Havering	1	0.287	0.384	0.317	0.233	0.066
Hillingdon	Kingston upon Thames	-1	0.171	0.384	0.333	0.258	0.222
Kingston upon Thames	Hillingdon	1	0.278	0.404	0.319	0.267	0.092
Kingston upon Thames	Lewisham	0	0.348	0.714	0.629	0.465	0.188
Lewisham	Kingston upon Thames	0	0.428	0.76	0.659	0.491	0.138

Table 33: *Coexplosive model, with explosive period 2007 - 2008*

a "ripple effect" across local London housing markets. Though it is worth noting, that there does not seem to be a dominant region in this process, or at least Central London as I aggregated did not act as a dominant region.

There are several explanations put forward for the causes of the ripple effect. One is that house prices reflect underlying information and expectations of the markets, which might not change immediately in all regions. Instead, information might transmit spatially, through nodes in a network. The evidence for this explanation is somewhat mixed from the first part of this chapter - if information does transmit spatially, it would seem natural to assume that it would come to the most central regions first. I.e. there would be a dominant region, where price changes occur in first, which then affects the surrounding regions. There isn't evidence that Central London acts as a dominant region, however. On the other hand, I would argue that the second part of this paper does support the explanation that the ripple effect is caused by the spatial transmission of information.

In the second part of this paper, I found evidence of coexplosivity between different regions of London. This is evidence of there having been bubbles in the house prices of these London boroughs, which were later transmitted to other London boroughs. Many of these coexplosive pairs occurred in regions that were spatially contiguous to each other. The main explanations of bubbles revolve around the information held and expectations of buyers in markets. Bubbles exist because of expectations of future price increases. Therefore, the spatio-temporal transmission of bubbles across regions would imply that the information was spatio-temporally transmitted as well.

Moreover, the fact that this occurred during a bubble excludes other explanations for the ripple effect. Two explanations for the ripple effect draw on the idea that it might be that there are other factors that have a spatial relationship, that then cause these spatial patterns in price changes. One explanation is that the actual ripple effect is in economic fundamentals - e.g. incomes increase in a ripple effect like manner, and then house prices increase because of this. However I specifically chose the timeframe here to exclude this - as I found in the first paper of this thesis that there were price increases in London housing markets over this period that could not be justified by the fundamentals. Another explanation for the ripple effect is that price changes are caused by migration - as one region becomes more expensive, people will move from this region into surrounding areas, increasing housing demand and prices. However, if this was what was driving house prices increases, we would expect to see this driving rental price increases as well. However, again the result of the first paper of this thesis is that house price increases were exceeding that justified by the fundamentals - which include rental prices. In other words, the price increases were greater than the rental price increases. Which one would not expect if price increases were due to increases in housing demand. So, overall, the results of this paper strongly support the narrative of the ripple effect being caused by the spatial transmission of information and expectations causing spatial effects in the changes of house prices.

V. CONCLUSION

In this thesis, I have presented three papers addressing topics in housing bubbles and housing economics. In the first paper, I spend much time addressing the issue of what a bubble is, and how it can be identified. The central theme of this is that when a bubble occurs, price increases are not proportionate to changes in fundamentals. With house prices, the key fundamental factors are i) the cost of housing services (which is determined by factors such as aggregate income and the available supply of housing), which can be measured by rental prices, ii) the availability of financing, in particular mortgage costs, and iii) the various costs associated with home ownership, such as maintenance costs. Currently, a popular method in the literature for detecting housing bubbles involves testing for explosiveness on price to rent or price to income ratios, in particular using the recursive unit root procedure developed by Phillips et al. (2011) and Phillips et al. (2015). The rationale is that house prices are determined by rental prices and incomes - if house prices are high relative to these factors, then this might be a sign of overpricing, and if this ratio has increased in an explosive manner, then this might be a sign of a bubble. However, these ratios ignore many factors, primarily the availability of financing. In the first paper of this thesis, I present a novel affordability ratio, the user cost to rent ratio, which captures all of the fundamental factors that drive house prices. I demonstrate how we can identify where housing becomes overpriced where this ratio exceeds one, and we can identify bubble periods by finding where changes in this ratio become explosive, indeed bubble periods are identified in the period around and before the Global Financial Crisis (GFC). Moreover, I show that using the user cost to rent ratio, one will not misidentify bubbles which would have been identified using the price to rent and price to income ratios. Using the price to rent and price to income ratios, a bubble is misidentified in the post-GFC period, where price increases can be explained by changes to interest rates over this period. While the user cost of housing is a difficult term to calculate, I show that these results are largely unchanged by the use of different risk premiums (which only affect the specific timing of under or overpricing), and are qualitatively largely unchanged by the use of different house price growth expectations.

The main conclusion from the first paper then, is to highlight the importance of interest rates in the determination of house prices. This is important for policy makers at the Bank of England and central banks elsewhere. A natural conclusion is that if lower interest rates can make house prices go up, surely now that interest rates have gone up house prices will come down? The simple answer to this question is that they already have gone down, particularly in London, and particularly in real terms[54], though not as much as the theory would suggest. Some simple calculations equating the user cost of housing to rental prices would suggest that London house prices would need to collapse by 25% to remain in equilibrium (and this is in nominal terms, before taking into account the effects of inflation).[55]. In other words,

[54]At the time of writing, according to the latest HPI data from the ONS, the average cost of a property in London decreased from a peak of £542,387 in August 2022 to £523,134 in June 2024: a nominal decrease of 3.5%. However, CPI also increased 8.9% over this period. Which means that in real terms, London house prices fell by 11.5%.

[55]A simple way to do this calculation is to assume that house prices are in equilibrium - in which case, the user cost of capital (the real interest rate term) for housing should be equal to rental prices divided by prices (the inverse of the price to rent ratio). In August 2022, average rental prices in London were £1773 per month, according to the Price Index of Private Rents, published by the ONS. To be conservative I shall use the average price for flats or maisonettes, which was £452,725 in London in August 2022. This would suggest a

house prices might not have been overpriced, but they are now, given that interest rates have gone up and housing prices have not declined proportionally. Does this mean there will be a housing market crash? As I discuss in the first paper, while this is a possibility, I do not think this is likely. If there is a lack of demand for houses, homeowners can simply not sell their property, and leave property vacant, or forestall moving, if they so wish. Homeowners are reluctant to sell property at a loss, so in a market downturn, homeowners will either not sell, or properties will remain on the market for longer. As I discuss in the first paper, transactions did indeed collapse in 2022, and have not recovered to their previous levels. There has been a particular decline in house purchases funded by mortgages. Figure 42 shows how new mortgage advances, gross and net of repayments, have collapsed since interest rates increased in 2022.

However, at the time of writing, there is a general expectation that interest rates will decline, with the Bank of England bank rate reaching around 3.5% by the end of 2025[56]: a decline of 1.5% from the current bank rate. The key issue is if the market can sustain current

user cost of capital for housing of 4.7%. Again, to be conservative I shall estimate this as 5%. According to statistics produced by the Financial Conduct Authority, the weighted average of fixed rate interest rates increased by approximately 3%, from 1.87% in Q1 2022 to 4.90% in Q1 2024. *Ceteris paribus*, this would mean an increase in the user cost of capital for housing from 5% to 8%, a 60% increase. Average rental prices increased from £1773 in August 2022 to £2114 per month July 2024, an almost 20% increase. Given a 60% increase in the user cost of capital, but only an increase of 20% in rents, house prices would need to fall 25% to remain in equilibrium. The calculation for England is extremely similar, with a user cost of capital at 5.3% in 2022 (from rent at £1135 monthly to prices of £258,215), with rents increasing to £1319 in 2024, giving an overall fall of 25.8% to remain in equilibrium.

[56] See, for instance, https://www.bbc.co.uk/news/articles/cldd6x6gglxo and https://www.forbes.com/uk/advisor/mortgages/mortgage-interest-rates-forecast/ accessed 11/09/2024

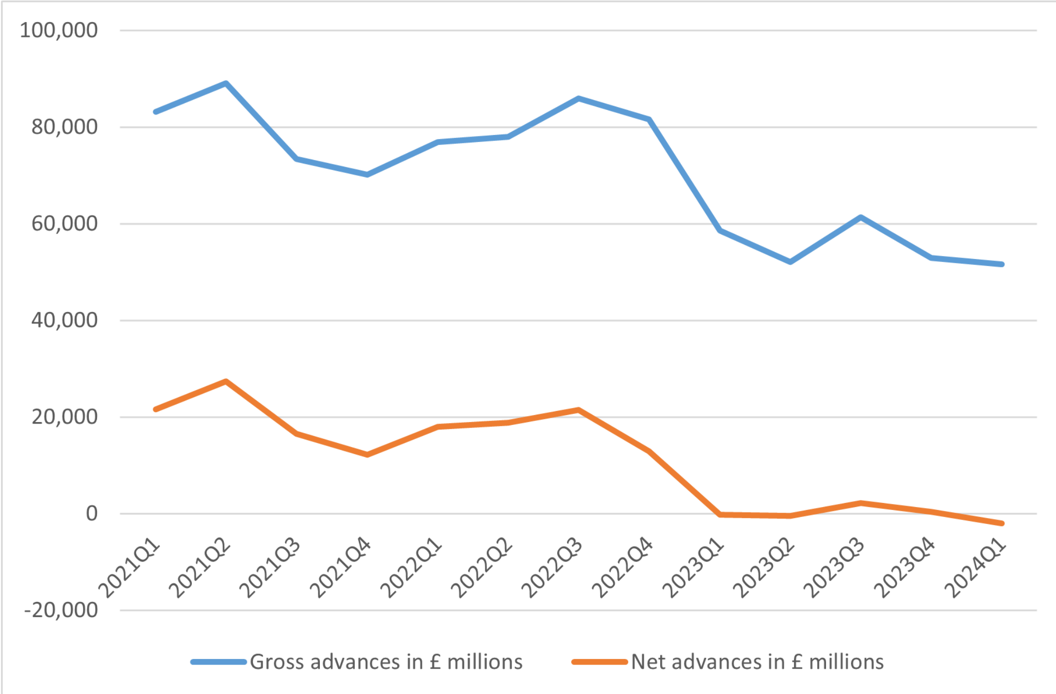


Figure 42: *Mortgage advances by quarter, 2021 to present*

levels of house prices until then. If there is a 2008-style crash, with declines in asset prices universally, then a housing crash would most likely also occur. However if there is not, it is possible that the housing market is kept afloat until interest rates decline. If mortgage rates also decline by 1.5% by the end of 2025, and if there are further increases in rental prices[57], then house prices might well not have to decline at all to remain in equilibrium - or at the very least, the decline would be in the single digits compared to the 2022 price level[58]. Of

[57]Which seems likely - according to the latest data published by the ONS, at the time of writing, rental price inflation is at 8.6% annually in the 12 months to July 2024, well above general levels of inflation.

[58]Based off of the calculation before - if the user cost of capital increases 30% instead of 60%, then overall rental price inflation would need to be only 30% for house prices to remain at the 2022 level. This would require only a further 10% increase in rental prices relative to 2022 between now and the end of 2025, and even if further rental price inflation was zero (which seems unlikely), this would entail a house price decline of only 8%.

course this is only if there is not a general economic crash. And if there is not a crash in the economy, then it is likely that inflationary pressures will remain high. It seems inevitable then, that if there is not a crash in nominal terms, there will have been a significant decline in real terms.

In the second paper of this thesis, I investigated whether the housing bubbles that I identified in London before and around the GFC were rational or irrational, using the coexplosive model of Nielsen (2010) and Engsted and Nielsen (2012). I found that the data overwhelmingly did not support the rational bubble model in the bubbles in London housing markets. This result provides evidence against the existence of rational bubbles, and suggests that bubbles form in an irrational manner. In the third paper of this thesis, I found evidence of ripple effects in London Housing boroughs, in particular price convergence and price dependence, using the spatio-temporal model of Holly et al. (2011). Finding ripple effects at the borough level in London is an interesting result, that is line with ripple effects being identified at other levels of the UK, and in other markets abroad. There are however many possible explanations for these ripple effects. I argue that when I also find evidence of coexplosivity, using the model of Evripidou et al. (2022), this is evidence that price changes occur due to spatio-temporal changes in expectations. This explanation posits that expectations change as information flows through spatial networks, and these changes in expectations change house prices. This would result in both the ripple effect, which there is evidence of in the data, and also bubble transmission between housing markets (coexplosivity), which is also seen in the data. On the other hand, other competing explanations for the ripple effect do not match the data: specifically, that the ripple effect is caused by spatial propagation of the fundamental factors of housing prices - I specifically chose the time-frame used because house prices are not proportionate to changes in other fundamentals. These two papers of this thesis highlight the importance of expectations in the determination of house prices and formation of housing bubbles. Again this result is important for policy makers, such as at the Bank of England and other central banks. Interest rates have been kept so low, for so long, there is a strong risk that this has affected perceptions and expectations of interest rates and house prices among the general public, which would undermine effective monetary policy.

To conclude, this thesis has taken a deep dive into issues not just in housing bubbles, but in housing economics more generally. During this research, some clear avenues for future research have been identified. In particular, the use of the market implied expected growth rate to identify bubbles, which I suggest in the first paper, is a prospect I think could prove fruitful in future. The expectations of price growth was by far the most difficult term to calculate in the first paper. However, by simply rearranging the formula, we can find the market implied expected growth rate - what the expected growth rate would need to be to justify current house prices given the current levels of rent, interest rates, etc. The key feature of asset price bubbles is that they are caused by the expectation of future price growth. Hence while this was not appropriate in the econometric framework I used in the first chapter, I suspect its use in a different econometric framework would prove highly useful in the identification of asset price bubbles in the future.

Bibliography

A. Abbott and G. De Vita. Pairwise convergence of district-level house prices in london. *Urban Studies*, 49(4):721–740, 2012.

M. Andrew and G. Meen. House price appreciation, transactions and structural change in the british housing market: A macroeconomic perspective. *Real Estate Economics*, 31 (1):99–116, 2003.

Y. Arslan, B. Kanık, and B. Köksal. Anticipated vs. unanticipated house price movements and transaction volume. *Journal of Housing Economics*, 28:121–129, 2015.

R. E. Bailey. *The economics of financial markets*. Cambridge University Press, 2005.

D. Baker. The housing bubble and the financial crisis. *Real-world economics review*, 46 (20):73–81, 2008.

M. Balcilar, A. Beyene, R. Gupta, and M. Seleteng. 'ripple' effects in south african house prices. *Urban Studies*, 50(5):876–894, 2013.

M. Bangura and C. L. Lee. House price diffusion of housing submarkets in greater sydney. *Housing Studies*, 35(6):1110–1141, 2020.

N. Barberis, A. Shleifer, and R. Vishny. A model of investor sentiment. *Journal of financial economics*, 49(3):307–343, 1998.

K. Barker. Review of housing supply: Delievering stability: Securing our future housing needs, final report. Technical report, London: HM Treasury, 2004.

T. Basse, M. Karmani, H. Rjiba, and C. Wegener. Does adhering to the principles of green finance matter for stock valuation? evidence from testing for (co-) explosiveness. *Energy Economics*, 123:106729, 2023.

J. Berkovec and J. G. Jr. Turnover as a measure of demand for existing homes. *Journal of Housing Economics*, 24(4):421–440, 1996.

B. S. Bernanke. Monetary policy and the housing bubble. 2010.

O. J. Blanchard and M. W. Watson. Bubbles, rational expectations and financial markets, 1982.

J. Bone and K. O'Reilly. No place called home: The causes and social consequences of the uk housing 'bubble'. *The British journal of sociology*, 61(2):231–255, 2010.

N. Braakman and S. McDonald. Housing subsidies and property prices: Evidence from england. *Regional Science and Urban Economics*, 80(1):103374, 2020.

G. Cameron, J. Muellbauer, and A. Murphy. Was there a british house price bubble? evidence from a regional panel. 2006.

J. Y. Campbell and R. J. Shiller. Cointegration and tests of present value models. *Journal of political economy*, 95(5):1062–1088, 1987.

J. Y. Campbell and R. J. Shiller. Interpreting cointegrated models. *Journal of Economic Dynamics and Control*, 12(2-3):505–522, 1988a.

J. Y. Campbell and R. J. Shiller. The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies*, 1(3):195–228, 1988b.

J. Y. Campbell, A. W. Lo, and A. C. MacKinlay. *The Econometrics of Financial Markets*. Princeton University Press, 1997.

J. Y. Campbell, A. W. Lo, A. C. MacKinlay, and R. F. Whitelaw. The econometrics of financial markets. *Macroeconomic Dynamics*, 2(4):559–562, 1998.

K. E. Case, J. M. Quigley, and R. J. Shiller. Comparing wealth effects: the stock market versus the housing market. *Topics in Macroeconomics*, 5(1):20121001, 2005.

J. Chen, Y. Chen, R. J. Hill, and P. Hu. The user cost of housing and the price-rent ratio in shanghai. *Regional Science and Urban Economics*, 92:103738, 2022.

P.-F. Chen, M.-S. Chien, and C.-C. Lee. Dynamic modeling of regional house price diffusion in taiwan. *Journal of Housing Economics*, 20(4):315–332, 2011.

M.-C. Chiang and I.-C. Tsai. Ripple effect and contagious effect in the us regional housing markets. *The Annals of Regional Science*, 56:55–82, 2016.

S.-h. Chiang. Housing markets in china and policy implications: Comovement or ripple effect. *China & World Economy*, 22(6):103–120, 2014.

G. L. Clark, R. Durán-Fernández, and K. Strauss. 'being in the market': the uk house-price bubble and the intended structure of individual pension investment portfolios. *Journal of Economic Geography*, 10(3):331–359, 2010.

S. P. Clark and T. D. Coggin. Was there a us house price bubble? an econometric analysis using national and regional panel data. *The Quarterly Review of Economics and Finance*, 51(2):189–200, 2011.

J. H. Cochrane. Presidential address: Discount rates. *The Journal of finance*, 66(4): 1047–1108, 2011.

S. Cook and D. Watson. A new perspective on the ripple effect in the uk housing market: Comovement, cyclical subsamples and alternative indices. *Urban Studies*, 53(14):3048–3062, 2016.

K. Daniel, D. Hirshleifer, and A. Subrahmanyam. Investor psychology and security market under-and overreactions. *the Journal of Finance*, 53(6):1839–1885, 1998.

E. R. de Wit, P. Englund, and M. K. Francke. Price and transaction volume in the dutch housing market. *Regional Science and Urban Economics*, 43(2):220–241, 2013.

B. T. Diba and H. I. Grossman. On the inception of rational bubbles. *The Quarterly Journal of Economics*, 102(3):697–700, 1987.

B. T. Diba and H. I. Grossman. The theory of rational bubbles in stock prices. *The Economic Journal*, 98(392):746–754, 1988a.

B. T. Diba and H. I. Grossman. Explosive rational bubbles in stock prices? *The American Economic Review*, 78(3):520–530, 1988b.

J. Dolinar. *Assessing the Canadian Housing Market: A User Cost Approach*. PhD thesis, University of Victoria, 2018.

J. A. Doornik. Approximations to the asymptotic distributions of cointegration tests. *Journal of Economic Surveys*, 12(5):573–593, 1998.

A. Dougherty and R. V. Order. Inflation, housing costs, and the consumer price index. *American Economic Review*, 72(1):154–164, 1982.

J. V. Duca, J. Muellbauer, and A. Murphy. What drives house price cycles? international experience and policy issues. *Journal of Economic Literature*, 59(3):773–864, 2021.

J. P. Elhorst et al. *Spatial econometrics: from cross-sectional data to spatial panels*, volume 479. Springer, 2014.

T. Engsted and B. Nielsen. Testing for rational bubbles in a coexplosive vector autoregression. *The Econometrics Journal*, 15(2):226–254, 2012.

T. Engsted and T. Q. Pedersen. Predicting returns and rent growth in the housing market using the rent-price ratio: Evidence from the oecd countries. *Journal of International Money and Finance*, 53:257–275, 2015.

T. Engsted, S. J. Hviid, and T. Q. Pedersen. Explosive bubbles in house prices? evidence from the oecd countries. *Journal of International Financial Markets, Institutions and Money*, 40:14–25, 2016.

G. W. Evans. Pitfalls in testing for explosive bubbles in asset prices. *The American Economic Review*, 81(4):922–930, 1991.

A. C. Evripidou, D. I. Harvey, S. J. Leybourne, and R. Sollis. Testing for co-explosive behaviour in financial time series. *Oxford Bulletin of Economics and Statistics*, 84(3): 624–650, 2022.

G. Fack. Are housing benefit an effective way to redistribute income? evidence from a natural experiment in france. *Labour Economics*, 13(6):747–771, 2006.

A. Finicelli. House price developments and fundamentals in the united states. *Bank of Italy occasional paper*, (7), 2007.

M. Flavin and T. Yamashita. Owner-occupied housing and the composition of the household portfolio. *American Economic Review*, 92(1):345–362, 2002.

J. Gallin. The long-run relationship between house prices and income: evidence from local housing markets. *Real Estate Economics*, 34(3):417–438, 2006.

J. Gallin. The long-run relationship between house prices and rents. *Real Estate Economics*, 36(4):635–658, 2008.

S. Gibbons and A. Manning. The incidence of uk housing benefit: Evidence from the 1990s reforms. *Journal of Public Economics*, 90(1):799–822, 2006.

E. L. Glaeser, J. Gyourko, and A. Saiz. Housing supply and housing bubbles. *Journal of Urban Economics*, 64(1):198–217, 2008.

R. Greenaway-McGrevy and P. C. Phillips. Hot property in new zealand: Empirical evidence of housing bubbles in the metropolitan centres. *New Zealand Economic Papers*, 50(1):88–113, 2016.

I. Grigoryeva and D. Ley. The price ripple effect in the vancouver housing market. *Urban Geography*, 40(8):1168–1190, 2019.

D. I. Harvey, S. J. Leybourne, R. Sollis, and A. R. Taylor. Tests for explosive financial bubbles in the presence of non-stationary volatility. *Journal of Empirical Finance*, 38: 548–574, 2016.

C. Hay. Good inflation, bad inflation: The housing boom, economic growth and the disaggregation of inflationary preferences in the uk and ireland. *The British journal of politics and international relations*, 11(3):461–478, 2009.

R. J. Hill and I. A. Syed. Hedonic price–rent ratios, user cost, and departures from equilibrium in the housing market. *Regional Science and Urban Economics*, 56:60–72, 2016.

C. Himmelberg, C. Mayer, and T. Sinai. Assessing high house prices: Bubbles, fundamentals and misperceptions. *Journal of Economic Perspectives*, 19(4):67–92, 2005.

D. Hirshleifer. Investor psychology and asset pricing. *The journal of Finance*, 56(4): 1533–1597, 2001.

S. Holly, M. H. Pesaran, and T. Yamagata. The spatial and temporal diffusion of house prices in the uk. *Journal of Urban Economics*, 69(1):2–23, 2011.

A. Holmans. *Historical Statistics of Housing in Britain*. Cambridge: Cambridge Center for Housing and Planning Research, 2005.

Y. Hu and L. Oxley. Bubbles in us regional house prices: evidence from house price–income ratios at the state level. *Applied Economics*, 50(29):3196–3229, 2018.

J. Huang and G. Q. Shen. Residential housing bubbles in hong kong: Identification and explanation based on gsadf test and dynamic probit model. *Journal of Property Research*, 34(2):108–128, 2017.

S. Johansen and A. R. Swensen. Testing exact rational expectations in cointegrated vector autoregressive models. *Journal of Econometrics*, 93(1):73–91, 1999.

S. Johansen and A. R. Swensen. More on testing exact rational expectations in cointegrated vector autoregressive models: Restricted constant and linear term. *The Econometrics Journal*, 7(2):389–397, 2004.

B. K. Kivedal. Testing for rational bubbles in the us housing market. *Journal of Macroeconomics*, 38:369–381, 2013.

D. Kulikauskas. The user cost of housing in the baltic states. *Journal of European Real Estate Research*, 2017.

H. H. Lean and R. Smyth. Regional house prices and the ripple effect in malaysia. *Urban Studies*, 50(5):895–922, 2013.

E. Levin, A. Montagnoli, and R. Wright. Demographic change and the housing market: Evidence from a comparison of scotland and england. *Urban Studies*, 46(1):27–43, 2009.

T. Loughran and J. R. Ritter. The new issues puzzle. *The Journal of finance*, 50(1):23–51, 1995.

G. Mankiw and D. Weil. The baby boom, the baby bust and the housing market. *Regional Science and Urban Economics*, 19(1):235–258, 1989.

A. Martin and J. Ventura. The macroeconomics of rational bubbles: a user's guide. *Annual Review of Economics*, 10:505–539, 2018.

M. Mayer. Are wall street's young millionaires paid enough? *New York*, 2(13):36–43, 1969.

G. Meen. The removal of mortgage market constraints and the implications for econometric modelling of uk house prices. *Oxford Bulletin of Economics and Statistics*, 52(1):1–23, 1990.

G. Meen. Spatial aggregation, spatial dependence and predictability in the uk housing market. *Housing studies*, 11(3):345–372, 1996.

G. Meen. Regional house prices and the ripple effect: a new interpretation. *Housing studies*, 14(6):733–753, 1999.

G. Meen and C. Whitehead. *Understanding Affordability: The Economics of Housing Markets*. Bristol University Press, 2020.

R. Meese and N. Wallace. Testing the present value relation for housing prices: Should i leave my house in san francisco? *Journal of Urban Economics*, 35(3):245–266, 1994.

V. Mikhed and P. Zemčík. Testing for bubbles in housing markets: A panel data approach. *The Journal of Real Estate Finance and Economics*, 38(4):366–386, 2009.

D. Miles. *Housing, financial markets and the wider economy*, volume 1. John Wiley & Sons Incorporated, 1994.

I. Mulheirn. Tackling the uk housing crisis: is supply the answer. *UK Collaborative Centre for Housing Evidence*, 2019.

B. Nielsen. Analysis of coexplosive processes. *Econometric Theory*, 26(3):882–915, 2010.

G. B. Northcraft and M. A. Neale. Experts, amateurs, and real estate: An anchoring-and-adjustment perspective on property pricing decisions. *Organizational behavior and human decision processes*, 39(1):84–97, 1987.

D. of Communities and L. Government. Fixing our broken housing market, 2017.

H. H. Pesaran and Y. Shin. Generalized impulse response analysis in linear multivariate models. *Economics letters*, 58(1):17–29, 1998.

P. Petris, G. Dotsis, and P. Alexakis. Bubble tests in the london housing market: A borough level analysis. *International Journal of Finance & Economics*, 27(1):1044–1063, 2022.

P. C. Phillips and S. Shi. Real time monitoring of asset markets: Bubbles and crises. In *Handbook of statistics*, volume 42, pages 61–80. Elsevier, 2020.

P. C. Phillips, Y. Wu, and J. Yu. Explosive behavior in the 1990s nasdaq: When did exuberance escalate asset values? *International economic review*, 52(1):201–226, 2011.

P. C. Phillips, S. Shi, and J. Yu. Testing for multiple bubbles: Limit theory of real-time detectors. *International Economic Review*, 56(4):1079–1134, 2015.

J. M. Poterba. Taxation and housing: Old questions, new answers, 1992.

I. Rherrad, J.-L. Bago, and M. Mokengoy. Real estate bubbles and contagion: new empirical evidence from canada. *New Zealand Economic Papers*, 55(1):38–51, 2021.

F. Schindler. Predictability and persistence of the price movementsof the s&p/case-shiller house price indices. *Journal of Real Estate Finance & Economics*, 46(1):44–90, 2013.

F. Schindler. Persistence and predictability in uk house price movements. *Journal of Real Estate Finance & Economics*, 48(1):132–163, 2014.

R. J. Shiller. Do stock prices move too much to be justified by subsequent changes in dividends? *The American Economic Review*, 71(3):421–436, 1981.

R. J. Shiller. *Irrational Exuberance*. Princeton University Press, 2000.

R. J. Shiller. From efficient markets theory to behavioral finance. *Journal of economic perspectives*, 17(1):83–104, 2003.

T. Sinai and N. S. Souleles. Owner-occupied housing as a hedge against rent risk. *The Quarterly Journal of Economics*, 120(2):763–789, 2005.

J. E. Stiglitz. Symposium on bubbles. *Journal of economic perspectives*, 4(2):13–18, 1990.

S. Susin. Rent vouchers and the price of low-income housing. *Journal of Public Economics*, 83(1):109–152, 2002.

R. Sustek. A back-of-the-envelope analysis of house prices: Czech republic, 2013-2021. *Working Paper*, 2021.

J. Tirole. Asset bubbles and overlapping generations. *Econometrica: Journal of the Econometric Society*, pages 1499–1528, 1985.

I.-C. Tsai. Ripple effect in house prices and trading volume in the uk housing market: New viewpoint and evidence. *Economic Modelling*, 40:68–75, 2014.

I.-C. Tsai. Information content of transaction volume: The housing market in the united kingdom. *International Journal of Strategic Property Management*, 22(5):348–357, 2018.

I.-C. Tsai and S.-H. Chiang. Exuberance and spillovers in housing markets: Evidence from first-and second-tier cities in china. *Regional Science and Urban Economics*, 77:75–86, 2019.

A. Tversky and D. Kahneman. Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *science*, 185(4157): 1124–1131, 1974.

A. Tversky and D. Kahneman. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292, 1979.

A. Tversky and D. Kahneman. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5:297–323, 1992.

W. Wheaton and G. Nechayev. The 1998-2005 housing "bubble" and the current "correction": What's different this time? *Journal of real Estate research*, 30(1):1–26, 2008.

C. Y. Yip, W. C. Wong, and H. E. Lim. Bubble detection in the malaysian housing market. *Malaysian Journal of Economic Studies*, 54(2):203–221, 2017.

X. Zhang, R. Kizys, C. Floros, K. Gkillas, and M. E. Wohar. Testing for rational bubbles in the uk housing market. *Applied Economics*, 53(8):962–975, 2021.

T. Zhi, Z. Li, Z. Jiang, L. Wei, and D. Sornette. Is there a housing bubble in china? *Emerging Markets Review*, 39:120–132, 2019.

W.-X. Zhou and D. Sornette. Is there a real-estate bubble in the us? *Physica A: Statistical Mechanics and its Applications*, 361(1):297–308, 2006.