UNIVERSITY COLLEGE LONDON

THE BARTLETT CENTRE FOR ADVANCED SPATIAL ANALYSIS

The spatial and temporal patterns of residential house prices and housing affordability in England

Bin Chi

Primary Supervisor: Dr Adam Dennett

Secondary Supervisor: Dr Thomas Oléron-Evans

In fulfilment of the requirements for the degree of Doctor of Philosophy

September 2021

Declaration

I, Bin Chi, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

____

Part of the work in Section 3.3 of Chapter 3 is published with the updated datasets in the publication:

> Chi B, Dennett A, Oléron-Evans T, et al. (2021) A new attribute-linked residential property price dataset for England and Wales, 2011–2019. *UCL Open: Environment*. 2021;(2):07. https://dx.doi.org/10.14324/111.444/ucloe.000019

The following publication is based on Chapter 4:

> Chi B, Dennett A, Oléron-Evans T, et al. (2020) Shedding new light on residential property price variation in England: A multi-scale exploration. *Environment and Planning B: Urban Analytics and City Science.* https://doi.org/10.1177/2399808320951212

The following publication is based on Chapter 5:

> Chi B, Dennett A, Oléron-Evans T, et al. (2021) Delineating the Spatio-Temporal Pattern of House Price Variation by Local Authority in England: 2009 to 2016. *Geographical Analysis*. https://doi.org/10.1111/gean.12287

I declare that the research for these publications was solely my own work and that I am the lead author of these publications. The contribution of the other three authors, Adam Dennett, Thomas Oléron-Evans and Robin Morphet, was purely editorial and advisory.

Abstract

Housing affordability is one of the most urgent issues facing the world. It has been a key political concern in the UK since the 1960s. The UK housing policy challenge is to reduce the gap between house prices and the housing people can afford. A more nuanced understanding of the housing affordability issues in England is essential to create a prosperous and equal country over the next century. Housing affordability is determined by two aspects, one is residential house price and the other is household income. The drivers behind the changes in the cost and affordability of housing are complex and operate at different scales. This research explores spatial and temporal patterns of housing affordability in England through an in-depth analysis of residential house price variations at small geographic levels.

This research overcomes the difficulty of understanding house price and housing affordability variation in England at small geographic scales where house price data and income data are imperfect and the process is complex. A comprehensive geo-referenced housing price database is constructed, along with a systematic analysis of the house price variation at multi-geographic levels and further separate at different time scales. Through modelling and visualisation we can gain a deeper understanding of the spatial and temporal variations in house prices. The following research specifically focusses on the local authority level with annual time categorizations. Then by combining and comparing house price variation at local authority level and household housing budget for different types of buyers, this research creates a new method for understanding housing affordability, while highlighting housing affordability spatial-temporal patterns in England at small geographic scales and for

different types of buyers. Suggestions regarding housing policy and planning are offered at the end in order to ease housing affordability issues in England.

Impact statement

This research takes a first step to look at residential housing cost and affordability at a micro-level from a geographical perspective. This piece of work contributes to the society in three main ways: firstly, the reusable data linkage methods along with the newly created spatial house price dataset offers a new research era in quantitative residential housing research in the UK. This newly created and updateable house price dataset not only offers greater flexibility for the exploration and understanding of house price variation over different scales for society, but is also able further expand to include dwelling energy efficiency to be upgraded in the drive to a net zero carbon economy. Secondly, the model-based descriptive approach used in this research is applicable to spatial-temporal house research for different spatial scales and time period focus. The series of multi-level models offer a new systematic reproducible research to better investigate and understanding house price variation across time and space. It is an extremely powerful for the exploration of big data. The third and most central contribution of this work is the newly designed house affordability index in terms of affordable property size for a given buyer. It offers a meaningful and comparable indications for society to better understand the housing affordability issues they face. The method not only enables individual buyers to understand their own affordable house size across space and time, but also guide policy makers and local authority with the information of affordable housing size to better deliver local residential housing. This prove useful to assist with promoting policies for a fairer society.

Acknowledgements

This thesis would not have been possible without the support and help of many people around me. First and foremost, I would like to extend my sincere gratitude to my supervisors, Dr Adam Dennett and Dr Thomas Oléron-Evans, not only for their insightful guidance but also their continuing support, encourage and comment. I am also thankful to my research mentor, Professor Robin Morphet, who is always offers a help and share his experience when I was struggling my research. I am very grateful to have such a good research mentor to in my PhD journey.

In addition, I also extend my deep appreciation to those people who have kindly assisted and significantly supported me during my PhD studies, including Rosie McKay, Yi Wang, Jane Galbraith (UCL Department of Statistical Science), Rob Liddiard (UCL Energy Institute), Sheng Zhou (Ordnance Survey), Martin Zaltz Austwick, Christopher Charlton (University of Bristol), Caroline Bray (Land Registry), David Lockett (Land Registry), Richard Milton (UCL CASA), Flora Roumpani, Zhoupeng Ren, Polly Hudson, Mengqiu Cao, Guanpeng (Gavin) Dong, Neal Hudson, Yuqi Liu, Cecilia Chan, Jessica Williamson (MHCLG), Jake Mulley (MHCLG), Chris Hutchinson (TravelTime), Anca Vlad (UK Data Service), Cristina Magder (UK Data Service), Alasdair Rae, Qiulin Ke (UCL The Bartlett School of Planning), Yu (Steven) Zhang (Billigence), Calvin Chung, Angela Cooper (UCL CLIE), Tony Mulhall (RICS) and Simon Rubinsohn (RICS). I would also like to thank all staff members and colleagues from the Bartlett Centre for Advance Spatial Analysis at UCL, they create a nice and friendly working environment for my research. All the errors remaining are my own.

I would also like to sincerely thank Ordnance Survey (OS) for supply of huge amount data in this research, in particular I would like thank Nick Groome, OS technique team and EDINA. I am grateful to UCL and Chinese Scholarship Council (CSC No. 201708060184) for funding my doctoral research since my second year.

Finally, I would like to thank my parents for their invaluable encouragement and unconditional love during my PhD studies at UCL.

Contents

List of Figures

List of Tables

Glossary of terms and abbreviations

Abbreviations

ABP	AddressBase Plus
CDRC	Consumer Data Research Centre
EHS	English Housing Survey
EPC	Energy Performance Certificate
FRS	Family Resource Survey
GCM	growth curve model
GFC	Global Financial Crisis
GTFS	General Transit Feed Specification
HPM	House price per square metre
ICC	intraclass correlation coefficient
LA	Local Authority
LCF	Living Cost and Food Survey
LR	Land Registry
LSOA	Lower Level Super Output Area
MLM	Multilevel Model
MMTL	MasterMap Topography Layer
MSOA	Middle Layer Super Output Area
NSPL	National Statistics Postcode Lookup
ONS	Office for National Statistics
OS	Ordnance Survey
SDLT	Stamp Duty Land Tax

PPD	Price Paid Data
RICS	Royal Institution of Chartered Surveyors
TOID	Topographic Identifier
TP	transaction price
UK	United Kingdom
UPRN	Unique Property Reference Number
VPC	variance partition coefficients

Terms

Housing budget the component of household income and/or other sources of capital used to secure accommodation

List of publications based on the thesis

Journal Articles

- **Chi B**, Dennett A, Oléron-Evans T, et al. (2021) A new attribute-linked residential property price dataset for England and Wales, 2011–2019. *UCL Open: Environment*. 2021;(2):07. https://dx.doi.org/10.14324/111.444/ucloe.000019

- **Chi B**, Dennett A, Oléron-Evans T, et al. (2021) Delineating the Spatio-Temporal Pattern of House Price Variation by Local Authority in England: 2009 to 2016. *Geographical Analysis*. https://doi.org/10.1111/gean.12287

- **Chi B**, Dennett A, Oléron-Evans T, et al. (2020) Shedding new light on residential property price variation in England: A multi-scale exploration. *Environment and Planning B: Urban Analytics and City Science.* https://doi.org/10.1177/2399808320951212

Data repository

- **Chi B**, Dennett A, Oléron-Evans T, et al. (2021) House Price per Square Metre in England and Wales. London Datastore. https://data.london.gov.uk/dataset/house-price-per-square-metre-in-england-and-wales

- **Chi B**, Dennett A, Oléron-Evans T, et al. (2021). A new attribute-linked residential property price dataset for England and Wales 2011-2019. Colchester, Essex: UK Data Service. https://reshare.ukdataservice.ac.uk/854240/

Working papers

Two working paper has been published in Journal article, which is not list below.

- **Chi B**, Dennett A, Morphet R, et al. (2020) Exploring local authority travel time to London effects on spatio-temporal pattern of local authority house prices variation in England. *CASA working paper 218*. https://www.ucl.ac.uk/bartlett/casa/publications/2020/apr/casa-working-paper-218

- **Chi B**, Dennett A, Oléron-Evans T, et al. (2019) Creating a new dataset to analyse house prices in England. *CASA working paper 213*. https://www.ucl.ac.uk/bartlett/casa/publications/2019/sep/casa-working-paper-213

Conferences

- GISRUK 2018: A method for representative house prices for small areas though Data linking Land Registry Price Paid Data and OS MasterMap
- GISRUK 2019: A new insight into residential house price variation across England through linking Land Registry Price Paid Data and Domestic Energy Performance Certificates
- GISRUK 2020: Delineating the spatio-temporal patterns of house price variation at local authority level in England
- GISRUK 2021: Understanding housing affordability to determine the best property search areas for homeowners moving out of London

Chapter 1 Introduction

1.1 Research background

Shelter is one of the most basic human needs (Maslow, 1943; McLeod, 2018), people need a home and want to live in a pleasant place. Nowadays, the majority of countries worldwide are facing a critical housing challenge (Tsenkova and French, 2011; UN-HABITAT, 2011, 2012). In developed countries such as the United Kingdom (UK) rising house prices since the mid-1990s have led to some problems (Knoll et al., 2017). Escalating housing prices reduce peoples' ability to buy or rent a dwelling. These housing difficulties are normally discussed under the heading of housing affordability (Hulchanski, 1995). Rising house prices relative to earnings continue to have a negative impact on housing affordability in the UK, especially in England. These housing affordability issues have been widely discussed in media and research communities (Barton and Wilson, 2018; Collinson, 2014; John, 2015; ONS, 2017e; Osborne, 2014). For example, the ONS housing affordability (ONS, 2017e: 2016) statistic shows that in 1997, houses in England and Wales were on average worth 3.6 times average earnings, but this had risen to 7.6 times in 2016. This continuously worsening housing affordability is caused by larger increases in house prices (259%) than the increase in earnings (68%). With house prices in some areas becoming prohibitively expensive, owning a house becomes more difficult for many low-to-middle income households, especially for younger groups or first time buyers (Alakeson, 2011; Clarke, 2015). This has not only led to 'generation rent' but has also resulted in the term 'Bank of Mum and Dad'(Coulter, 2017, 2011) – a term used to describe the fact that many young adults rely on their parents for financial help to purchase their property (Cosslett, 2017; Doward, 2016). Those people unable to buy are renting for a longer period of time. Increasing demands in the rental sector also push up the cost of renting (Kollewe, 2017). Rising rental prices result in households spending an increasing proportion of their income on rent, which possibly leads to a lower quality of life (Ahmed, 2017; Laura and Vidhya, 2014). At the same time landlords obtain greater profits. This increase in the wealth gap between the house owners and everyone else can then aggravate the housing inequality.

There is both a rising cost of housing and a widening in wealth inequality in the UK, especially in England (Dorling, 2014; Levin and Pryce, 2011). England is the main contributor to the increase in UK house prices based in the ONS UK price index (ONS, 2017g). For the time period between January 2009 and January 2016, the English house price index rose from £163,000 to £220,000, indicating a 35% increase. Meanwhile, the house price index in the other three countries (Scotland, Wales, Northern Ireland) of the UK were mainly below £160,000 with a varying but relatively small price change (+/- 10%) for the same period. There is no doubt that England is facing a continuing critical challenge in access affordable housing, especially in some expensive areas (Department for Communities and Local Government, 2017; Hudson, 2018). Housing affordability issues result from the interplay between residential house price (the costs of renting or owning a house) and a household's available "housing budget" (Mulliner et al., 2016; Whitehead, 1991). The household's available "housing budget" is defined for this research as the component of household income and/or other sources of capital used to secure accommodation. The drivers behind the changes in the cost and the different households' housing affordability are complex and operate at different scales, from the macro political, economic and demographic drivers to the local dynamics of redevelopment, gentrification and evolving household characteristics (Smith, 1987). However, this spatial heterogeneity in house prices and housing affordability is normally crudely expressed at regional or local authority (LA) level (ONS, 2012a, 2017e; Shelter, 2015). To date, little systematic quantitative analysis has been conducted to unpick the spatial and temporal variations in cost and affordability of housing across England, particularly for small areas (ONS, 2015b).

Deficiencies in residential house price data hinder research on house price and housing affordability in England, especially for small areas. There is no official full coverage rental price dataset in England. Zoopla data are the most commonly used rental price data in England, but they have very a low number of records in some areas within England and the ONS argues that Zoopla is weak in reflecting the whole picture of the rental housing market (ONS, 2018c). However, the Land Registry Price Paid Data (LR PPD) shows a better geographic coverage than the Zoopla data. LR PPD is the administrative dataset from Her Majesty's LR. This official transaction price (TP) dataset is able to support statistics on small areas and offer a fuller picture of residential price in the owner-occupier market (South and Henretty, 2017). Given the low quality of available rental price datasets, this research only focusses on residential house price in the owner-occupier market.

Despite the importance of housing affordability in England, the current understanding of housing affordability is often limited by crude measurements resulting in imprecise interpretations. The UK's current housing policy challenge is to reduce the gap between house prices and the housing people can afford, in order to allow more people to own their home (Department for Communities and Local Government, 2017). A comprehensive understanding of house price and housing affordability in the owner-occupied housing market will assist policymakers in offering tailored and effective solutions at a local scale. Housing affordability is determined by two aspects, one is residential house price and the other is a household's housing budget. The main factors resulting in changes in the cost and affordability of housing are multifaceted. Furthermore, these factors have varying influence on the cost and affordability of housing at different scales. A more nuanced understanding of housing affordability issues, through in-depth analysis of residential house price variation, at small geographic levels, will more effectively support the development of useful polices to create a prosperous and more equitable Britain over the next century. This context shapes the research aim of this thesis.

1.2 Research aim and objectives

Considering that each of the four countries of the UK have differing housing policies and related legal frameworks (Best, 1996) and that England faces arguably a more substantial housing affordability issue than the other three countries, this research only focuses on England. This research aims to explore the geography of housing affordability in England, through an in-depth analysis of residential house price variations, at small geographic levels. In order to achieve this aim, a number of focused research objectives are proposed:

1. To investigate substantive literature and the current methodological techniques on house price variance and housing affordability with a more specific focus on the UK context.

2. To examine and review house price datasets and income datasets in England from public open datasets and identify the data deficiencies in understanding the house price

variation in order to create methods to overcome data deficiencies.

3. To build on prior methods and develop a reusable research framework to explore the housing variation at multiple scales and choose an appropriate house price indicator for the given geographical level.

4. To build on the research findings and further explore temporal house price variation.

5. To consider the spatial and temporal pattern of house price variation and to develop an effective method to reflect spatial-temporal housing affordability for different types of buyers.

6. To offer specific recommendations on current UK housing policy and planning policy.

1.3 Thesis Structure

This thesis is divided into seven chapters. The first three chapters provide the research background, research aims, a literature review and a data review. The following three empirical chapters separately explore house price variations at four geographic scales, house price trends at LA level and housing affordability at LA levels in England between 2009 and 2016. The last chapter concludes with a summary of the main findings of this research, as well as housing policy and planning policy suggestions, discusses the limitations, and proposes an agenda for further research. The detailed contents of the chapters are as follows:

Following the Introduction, Chapter 2 reviews the existing literature on house price and housing affordability and focuses particularly on four aspects, namely: house price in the UK, house price and housing affordability, the housing affordability debate and housing dynamic research. After systematically reviewing the existing literature, the chapter ends with research gaps and proposes the research question.

Chapter 3 reviews the available residential house price data and income data that are currently used in understanding the house price and housing affordability patterns in England. It starts with exploring available residential house price data in terms of quality, coverage and accuracy. The most comprehensive house price dataset is chosen to assist with answering the first research question: to what extent does residential house price vary at small geographic levels? Deficiencies found in the house price data are addressed and overcome by building up a comprehensive spatial house price database. Two data linkage methods are created to overcome the data deficiencies. The match rate, one statistics test (Kolmogorov–Smirnov test) and one differences measure (J-divergence) are used to identify the transaction information that is lost after the data linkages. Based on the amount of lost information, this chapter identifies the most appropriate period for the research. Finally, this chapter reviews the available income data to assist in the exploration of housing affordability issues in England.

Chapter 4 is the first analytical chapter. It focuses on understanding the house price variation at different geographical scales. Four-level variance component models are used to support a comprehensive and systematic analysis of the mean house price and house price variance at four multi-geographical scales (LA, MSOA, LSOA and individual transaction level). Two housing price measures (TP and house price per square metre-HPM) are selected for comparison, to investigate which is the better indicator to represent house price variation patterns in England.

Chapter 5 is the second analytical chapter. It is based on the results from Chapter 4 and continues to further investigate the house price trend at the proper geographical scales (i.e. LA and MSOA level). With a control of the proper geographical scales identified in Chapter 4, this chapter starts by exploring three different time effects (yearly, half-yearly and quarterly) on house price spatio-temporal variation. Growth curve model (GCM) is used to offer a model-based description of the house price variation across different space, and time scales. Three independent GCMs are built to investigate the three time effects on house price variation one-by-one. Since using the yearly time scale fits the model best, LA annual house price trajectories are explored in the following analysis. Based on the LA's house price trajectory in terms of starting-price and overall house price change, hybrid hierarchical k-means clustering and choropleth mapping are used at the end to unlock the spatial and temporal patterns of the LA level house price in England.

Chapter 6 is the third analytical chapter. This chapter contains a three-stage workflow to create a new housing affordability metric based on the newly created house price data (HPM) and English Housing Survey (EHS). This new approach considers both the house price variation by property type and households with different housing budgets. It starts with the determination of whether it is necessary to consider house price variation by property types at LA level. The chapter then defines three typical household scenarios (cash buyers, mortgage buyers and home movers) with a further eight typical households with different housing budgets. Through the combination of the information of the above two stages, a new housing affordability proxy, in terms of affordable property size varying with property type, is created. This reveals the housing affordability patterns in England underlying different housing budgets of buyers at LA level and the related change of housing affordability patterns across space and time. Based on this new housing affordability measure, housing affordability patterns and trends at LA level in England are represented at the end of the chapter.

Chapter 7 synthesises the overall findings of the research. Through the combination of the new insights into housing affordability, categorised by different typical housing buyers and the spatial and temporal pattern of house price variation, at and below LA level, specific recommendations are given for future research. These recommendations are relevant to local and national housing policies and ongoing housing supply strategies. The chapter concludes the thesis with comments on the limitations of this research and with recommendations for future studies in this field.

Chapter 2 Literature review: house price and housing affordability

2.1 House prices in the UK

2.1.1 The evolution of national houses price in the UK

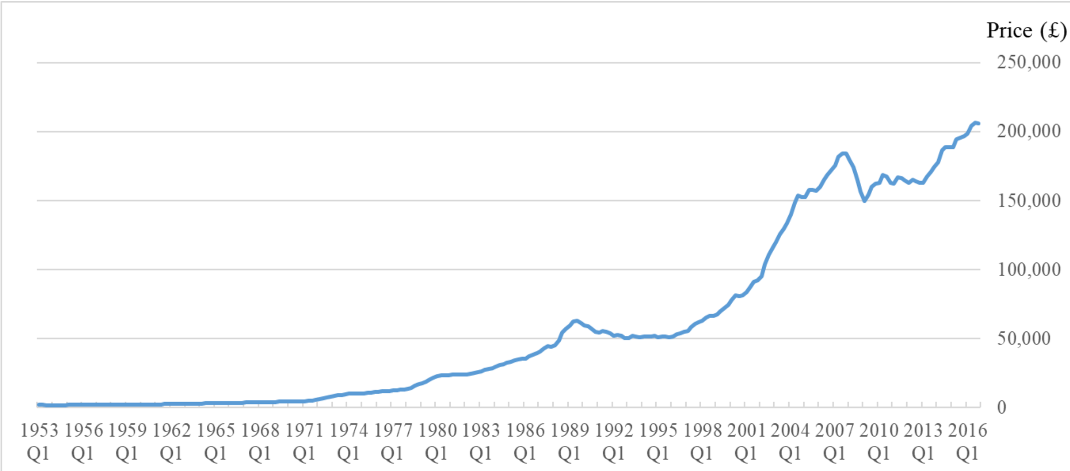


Figure 2.1 House price change in UK (1953-2016)

House prices in the UK have significantly increased over the past 60 years. Figure 2.1 House price change in UK (1953-2016) shows the house price trend between 1953 and 2016, based on the house price index dataset from the Nationwide Building Society (Nationwide Building Society, 2019). The overall UK house price increased relatively slowly from 1953 and reached its first peak in 1988, which was followed by a sharp drop until 1995. After this, house prices soared until the 2007 economic crisis. After the biggest fall ever recorded between 2007 and 2009, house prices started to rise again. According to the data in Figure 2.1, house prices at a national level increased by 30% between 1953 Q1 and 1990 Q1, while it rose by a historically unprecedented 244% between 1995 Q1 and 2007 Q1. Between 2009 Q4 to 2016 Q4 the average rate of increase was 33%.

2.2.2 The history of UK housing policy after World War II

The UK housing crisis has been developing since the 1960s (Lund, 2017). Government policies in response to this housing crisis are highly political and influence the housing market movement (Aha et al., 2018). Housing shortage is one of the main driving factors (Stephens, 2012; Swank et al., 2003). The population in the UK (Figure 2.2) consistently shows an increasing trend and grows rapidly in two periods, one is during the 1960s and the other is in the late 1980s. The year 1979 was particularly significant for the UK housing system in the period since World War II. Before 1979, the government focused on building more dwellings in order to increase the new housing supply. As shown in Figure 2.2, the newly built dwelling completions increased until the 1970's with a peak in 1968, after which there was a significant decline in dwellings completions, as there was a reduction in completions delivered by the LA sector. After 1979, increasing home-ownership was treated as a key element of government housing policy. This is because the government recognised that homeowners formed a larger proportion of voters, and therefore housing-related policies could influence their voting behaviour. The changing of UK housing-related policies with highly political motives shaped the UK's current housing crisis. For example, Margaret Thatcher's government introduced the 'Right to Buy' to help people living in council properties into home ownership (Lund, 2017; Millins and Murie, 2006). This housing policy transferred more than 2.85 million social houses into private ownership, between 1980 and 2015 (Murie, 2016). After 1997, new housing policies such as "Help to Buy" continued focusing on helping people get on the property ladder (Dorling, 2014). As the ratio of supply to demand shrank after the 1970s, house price in the UK (Figure 2.1) rose rapidly. The Government's recent housing White Paper (DCLG, 2017) states that the UK housing market is "broken" mainly because of its failure to deliver enough affordable housing in appropriate locations over the long term.

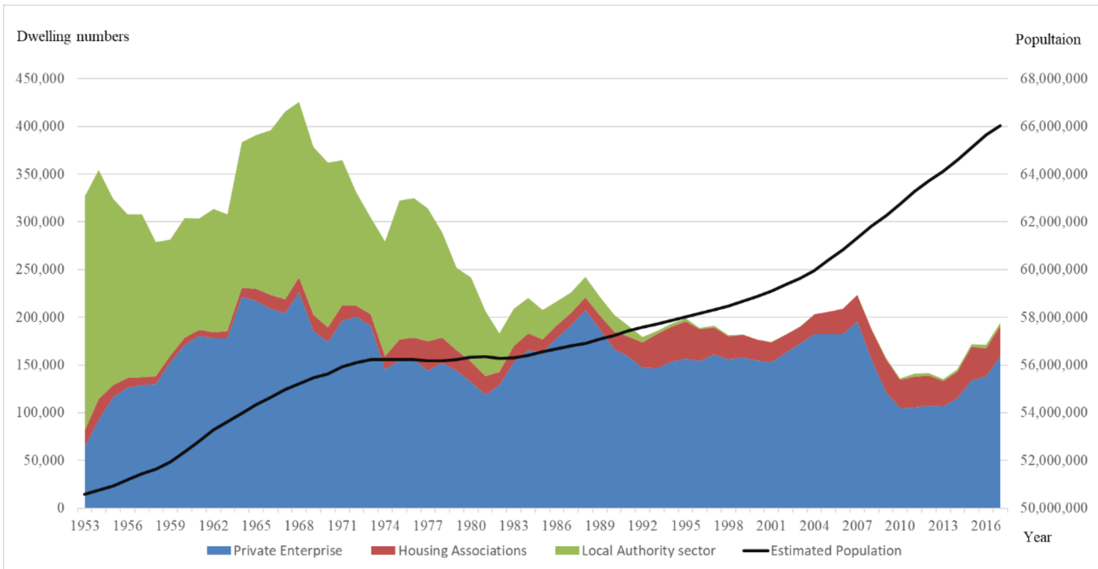


Figure 2.2 UK annual new build dwellings completed and estimated population (1953-2016)[1]

Alongside the changing housing policy and increasing population after World War II, housing tenure patterns in the UK have greatly shifted over the last century. The housing market in the UK includes owner occupiers, social renters and private renters, these are the three major categories in housing tenure. Owner occupiers include those who own outright and also those who are buying with a mortgage. The social renters include people that rent form LA or housing associations. Before the 1960s, the majority of people lived in the private rented sector while after that owner occupation grew to become the dominant type of tenure. In 2016, nearly two thirds (63%) of households lived in their own houses (owned either out-right). The remaining one third of household are split nearly equally between private and social renter sectors. However, the current UK housing tenure system is experiencing a new turning point,

[1] Resources: DCLG live table 241, ONS UK population estimated 1951-2014, UK population estimates and projections, 1960 to 2030.

in which the owner-occupied sector has declined and the private rental sector has increased significantly. After the period 2000 to 2016, the private rented sector increased from 10% to 20% of total housing, whilst the proportion of owner occupiers and social renters both showed a declining trend.

Currently, the UK's housing market is facing a continuous and chronic housing shortage. The ideology of home ownership which caters to the vast majority of the UK's citizens housing preferences plays an important role in this shortage of housing (Hilber and Schöni, 2016; Whitehead and Williams, 2011). Home ownership attainment is the prevailing force in housing policies since the 1980s, and imperceptibly influences the nation's housing market. According to the British Social Attitudes Survey, the vast majority (84%) of people in the UK would choose to buy accommodation rather than rent and this aspiration has remained broadly stable since 1996, with a slightly increasing trend (MHCLG, 2019). There is no doubt that recently demand-side housing policies such as Help-to-Buy may be popular among voters, but these policies fail to tackle the root causes of housing shortages and housing affordability (Department for Communities and Local Government, 2017; Hiber, 2013; Hilber and Schöni, 2016).

2.2 Housing costs and affordability

Housing unaffordability is at the heart of the housing crisis in England. England is the most expensive country in the UK and is becoming prohibitively expensive in some areas (Edwards, 2016; ONS, 2017). These escalating housing prices reduce people's ability to buy houses in England. Obviously, this has more of an influence on people with a low housing budget, such as low-to-middle income people, poorer and younger groups. Soaring house prices have not only led to 'generation rent' but have also resulted in the phenomena of 'Bank of Mum and Dad', which means that as a result of prohibitively high house prices many young adults need to rely on their parent's financial help to purchase their first property (Cosslett, 2017; Doward, 2016).This, of course, means that access to housing then becomes contingent on the wealth and generosity of family. Those who do not have access to these resources are then systemically disadvantaged. For the people unable to buy, they are renting for longer. Increasing demand in the rental sector also pushes up the cost of renting (Kollewe, 2017). These rising rental prices mean that households spend an increasing proportion of their income on rent, thus possibly leading to a lower quality of life (Laura and Vidhya, 2014; Ahmed, 2017) while landlords obtain greater profits. This increase in the wealth gap between the house owners and everyone else contributes to social inequality and social immobility.

One dimension of housing affordability relates to the cost of property. The dynamics of the housing market significantly affect different households and generations (Lamont and Stein, 1999). For better-off households, most housing is affordable. While for others, no housing is affordable in an open housing market. So understanding house price dynamics is essential for understanding the housing affordability issue. Besides, understanding the change of house price is important for other reasons. First, housing is one of the most important components of household wealth (Di, 2001). At the household level, the change of house price affects household consumption behaviour and wellbeing. For example, if house prices increase, the house owner may re-mortgage and so get more money from the lender and fund increased consumption (Reinold, 2011). Meanwhile, a rise in house price for the potential buyers could delay the potential buyers' plan to buy a house. When an economic downturn is present, the income of potential buyers could decrease and homeowners may be forced to sell their homes if they are unable to afford their mortgage repayments, this can put the banking system at risk. Second, at the aggregate level, house price and the changes of house price have potential impact on the rest of the economy (Miller *et al.*, 2011; Pryce *et al.*, 2011). For example, when the house price is too high that no profit will gain from new construction, this will adversely affect the local economy. Collapses in house prices can cause financial crises (Iacoviello and Neri, 2010). Thus, the booms and busts in housing markets have been an issue of concern for policy makers. In addition, the transaction costs of house purchase contributes to the local economy (Pryce et al., 2011), these can include anything from estate agent, legal or surveyor fees. Therefore, an understanding of the value of houses is vital for decision making by individuals and by local and national government.

2.3 The housing affordability debate

Although housing affordability has been widely discussed in media and research communities for decades (Bogdon and Can, 1997; Bramley, 1994; Burke and Ralston, 2004; Fingleton et al., 2019), the measurement and definition of it remains a challenge. The ratio of house price to income approach and the residual income approach are general measurement of housing affordability in the literature (Hulchanski, 1995; Stone et al., 2011). The residual income approach is based on the normative stand of non-housing expenditures left after paying for housing costs (Stone, 2006a, 2006b). The residual income approach aims to estimate whether the household could meet a basic non-housing consumption level after paying for housing, which has been widely used in the US. It considers the basic level of consumption for different household compositions or types. In any practical application, non-housing cost for the same disposable income level not only varies with different size and type of household but also change over space and time, thus the residual approach is not flexible in being universally applied. The ratio approach is based on the idea that housing cost should not exceed a certain fraction of household income (Fingleton et al., 2019). This indicator often uses median house price to median income. The ratio approach is the most common measurement used in the UK to assess housing affordability. All the current published ONS housing affordability index data use this ratio approach with a range of income data resources. Initially, ONS quantifies the housing affordability at three geographical scales (i.e. countries, regions and LAs) in England, through calculating the ratio of median house price to median gross full-time annual salary (John, 2015; Meen, 2018). There are two ways to calculate this affordability ratio using two different earnings approaches, either workplace-based earnings or residence-based earnings. The ONS therefore estimates the average income for an area based on either those who live there, or on those who work there. The temporal coverage of the workplace-based earnings approach is longer than the residence-based approach. Hence the workplace-based approach is more often used by ONS. Full coverage of workplace-based household income data in England are not available, so the earning information in ONS housing affordability estimation is partly based on the Annual Survey of Hours and Earnings (ASHE). The housing affordability ratio index was first published at Middle Layer Super Output areas (MSOAs) level in 2020, prior to which it had only been published at LA level. This was also the first time it was categorised by property type (Detached, Semi-Detached, Terraced, Flats/Maisonettes) (ONS, 2020a). Meanwhile, median gross full-time annual salary is replaced by net annual household income (equivalised) before housing costs, which is obtained from the Family Resources Survey (ONS, 2020b).

There is no official "rule of thumb" standard ratio in the UK, previous research tends to set the threshold of affordable housing costs at between 30% and 35% of net household income (Kutty, 2005). The UK housing charity Shelter classified housing as unaffordable if housing payments cost more than 35% of household net income, after tax and benefits on housing (John, 2015). The Cambridge Centre for Housing and Planning defines housing as affordable when the house price is 3.5 times the annual gross income for a single household or 2.9 times for a dual-income household (Whitehead et al., 2008). In the United States, the simple "rule of thumb" ratio standard has been 30% since the 1980s (Hulchanski, 1995). Housing practitioners generally agreed that housing is affordable if housing costs occupy less than 30 percent of tenants' income (Joice, 2014). Moreover, the Annual Demographia International Housing Affordability Survey used median house price divided by the median household income to measure housing affordability. Then they sort the ratio into four different affordable categories (Table 2.1) to compare major metropolitan housing markets in Australia, Canada, Hong Kong, Ireland, Japan, New Zealand, Singapore, the UK and the United States (Demographia International Housing Affordability Survey, 2018).

Table 2.1 Demographia International Housing Affordability Survey Affordability Ratings[2]

Rating	Median Multiple
Severely Unaffordable	5.1 & Over
Seriously Unaffordable	4.1- 5.0
Moderately Unaffordable	3.1- 4.0
Affordable	3.0 & Under

[2] Sources: http://www.demographia.com/

The traditional affordability ratio measurement is easy to construct as it only concentrates on averages. Consequently, this is the most common approach in the UK and gives limited information on how affordability varies in relation to varying household income and house prices (ONS, 2020c). The population of a country, county or city consists of different households in different locations with different social status and different income levels. Stone (2006) argued that the standard of affordability should also consider to whom the affordability applies. In the UK context, the fraction of household income spent on housing costs varies in different tenures. According to the ONS family spending report on weekly household expenditure (2016) across the UK, for rented households, 36% of income was spent on housing, whilst for owner occupied households, this fraction changed to 34%, and this proportion decreased to 30% for the mortgaged households (ONS, 2017a). This suggests that the ratio is too weak to accurately inform policy (Meen, 2018).

Recently, alternative measures for a housing affordability ratio approach continue to be developed to reflect housing affordability for different types of buyers. They are able to capture a certain household group which is assumed to be suffering serious affordability issues or capture widely different household circumstances. For example, two first-time buyer housing affordability indicators were created by ONS in 2018 to reflect first-time buyers' affordability (ONS, 2018a). The first approach calculates the ratio of lower quartile house prices to median gross annual workplace-based earnings for full-time workers aged 22 to 29 years. It is based on the assumption that first-time buyers, with average ages of around 30 years, are likely to purchase properties towards the lower end of the house price range. The second approach reflects the first-time buyer experience using mortgage data. This method uses the ratio of median house price purchased by first-time buyers with a mortgage to the median gross annual income of the mortgage applicants. This second approach is achieved using mortgage data from UK Finance, which is not openly accessible for academic analysis. However, these two approaches were conducted at two different geographic scales and are not comparable. The former is at regional level while the latter is at LA level. These two new approaches were only used once and then replaced by another package of affordability indicators in 2020 (ONS, 2020c). This new package of indices covers five different aspects of affordability (Table 2.2), which more closely reflect housing affordability reality. The five indicators are constructed by dividing each house price decile/quartile, by each income decile/quartile. Within this new package, only the purchase affordability by property type is available for small areas. In practical applications, there is an increased interest in tracking the housing affordability for a certain group of people, such as first-time buyers and low-income households (ONS, 2018a; Shelter, 2015; Easton, 2013).

Table 2.2 The latest alternative ONS housing affordability measurements in England

Tenure type	**Definition**	**Measurement**	**Geographical scales**	**Dataset**
Owner occupied	Purchase affordability	Ratio of median house prices by income deciles	Country and region	House price statistics for small areas, Living Costs and Food Survey
	Upfront costs involved with purchasing a residential property	10% deposit size plus Stamp Duty amount for 50th percentile of the median house prices by Lower layer Super Output Areas within a region.	Country and region	House price statistics for small areas, Living Costs and Food Survey, HM Revenue and Customs – Stamp Duty Land Tax rates

Tenure type	Definition	Measurement	Geographical scales	Dataset
	Mortgage repayment affordability	Proportion of household income spend on monthly mortgage repayments, by income deciles (assumption based on 10% deposit with 25-year mortgage term and fixed two-year mortgages interest rates)	Country and region	House price statistics for small areas, Living Costs and Food Survey, Interest and exchange rates data (Bank of England)
	Purchase affordability for small areas	Median house prices by net household income (annualised mean equivalised before housing costs)	MSOA	House price statistics for Small Areas, Income estimates for small areas
Private rented	Private rental affordability	Proportion of renter household income spend on rent, by income quartile	Country and region	Family Resources Survey, Private rental market statistics (VOA)

The latest alternative ONS housing affordability measurements (Table 2.2) achieve a greater degree of success in terms of reflecting housing affordability in relation to varying house prices and household incomes. They are available to calculate over time, but still have some disadvantages in common with their predecessors. First, the alternative affordability is presented either at regional level or at MSOA level in England. The affordability at LA level remains blank. Second, these ratio approaches are unable to directly reflect the extent of households' affordability. The term "affordable" is general considered to mean that housing costs make up less than 30% of household income. This 30% is a 'rule of thumb' used in quantifying housing affordability. The relationship, however, between household income and expenditure on housing can vary. For example, high-income households can more comfortably spend a higher percentage of their income on housing (Meen, 2018).Third, the majority of these ratio approaches implicitly assume all household within a survey want to enter the housing market. They fail to distinguish the households who do and who do not get on the housing ladder. Fourth, the ratio approach reflects the household affordability within a given area and is thus unable to reflect the housing affordability change for the same household moving between areas. Such information is an essential consideration for first time buyers or households seeking to move home.

To overcome some of the disadvantages of the affordability ratio approach, there is a growing scholarly interest in developing an optimal measure of housing affordability (Ezennia and Hoskara, 2019). Within the UK, the BBC's online housing calculator "Where can I afford to rent or buy?" achieves a degree of success in terms of integrating property size into the concept of affordability (Bailey et al., 2020). It presents an interactive platform allowing users to set their desired number of bedrooms, monthly payment amounts and available deposit, then automatically estimates housing affordability at LA level for either purchase or rental scenarios. This more detailed approach to the question of housing affordability offers a clear advantage over ratio-based methods, with evident practical value to potential home buyers. The BBC's online housing calculator, for the first time, offers the freedom to consider the household's housing budget difference in the housing affordability measure. However, compared with the latest alternative ONS housing affordability measurement (Table 2.2), the BBC approach fails to consider house price by property type.

The research above offers a picture of housing affordability at different geographic levels in a given period, but the results cannot be directly compared with each other in terms of different measurements. Housing unaffordability is at the heart of the housing crisis in England. Although the definitions for housing affordability are varied (Ezennia and Hoskara, 2019; Fingleton et al., 2019), there is a recognition that housing affordability measures aim to reflect detailed housing stress information and assist policymakers. The current understanding of housing affordability in England comes mainly through the house price to earnings/income ratio approach, such as the ONS housing affordability ratio. Since this ratio approach only concentrates on averages, and it is difficult to create targeted housing policies though its use, new measures of affordability are proposed by considering the available datasets (John, 2015; Meen, 2018; ONS, 2020c). Some of the new measures consider the house price difference according to property type, while some other measures consider different household circumstances. There is no doubt that England needs a housing affordability indicator that offers detailed insight into the dynamics of housing affordability and is applicable to different households (e.g. first-time buyers). A new approach can be created by emulating and combining the strengths of existing indicators. For example, through emulation and combination of methods similar to the BBC calculator and alternative ONS housing affordability measurements, to reflect housing affordability interplaying with price, property type, household circumstances (e.g. first-time buyers or mortgage buyers) and location.

2.4 House price dynamics research in the UK

Currently, the two principal aspects that are used to measure housing affordability are residential house price and a household's available "housing budget". The household's available "housing budget" is defined for this research as the component of household income and/or other sources of capital used to secure accommodation. The dynamics of the house price affect different households and generations (Lamont and Stein, 1999). House price will vary relatively across space and time – in different places and at different times, housing will be more or less affordable to different groups. So, understanding house price dynamics is essential for understanding housing affordability issues.

2.4.1 House price variation at macro geographic scale

Since the 1980s, the spatial diffusion of regional house prices has been a popular area in housing research in the UK (Drake, 1995). This spatial difference in house price change is normally conceptualized as a ripple effect (Cooper et al., 2013), which refers to the notion that house price shocks in one regions affect house price in other regions during a certain time period (MacDonald and Taylor, 1993). Empirical works exploring house price changes among regions in the UK show the same ripple-type pattern, which is London and the South East playing a leading role in terms of spill overs to other regions (Alexander and Barrow, 1994; Cook and Watson, 2016; Giussani and Hadjimatheou, 1991; MacDonald and Taylor, 1993). At the same time, Meen's (1999) research showed that this pattern not only exists in the long time frames, but also in the short time frames. Holly et al. (2011) used the average geometrical distances from London to particular towns/cities in each region as a proxy of the commuting distances from these various regions to London. This study not only found that the London area is a causal factor for house price increases of all regions, but also found that the closer the region is to London the more rapid an interaction exists. Additionally, houses price dynamics were examined in terms of housing type, such as property vintage (Cook, 2006; Cook and Holly, 2000; Gray, 2015; Narayan and Narayan, 2011) and property types (Morley and Thomas, 2016). This will impact upon the ripple effect. The majority of recent research on ripple effects in house prices uses aggregate housing without differentiating between house types. Only a few studies have begun to examine whether there are differences in the existence of a ripple effect for different housing types. Chris Hudson (2018) explored the interactive relationship among house prices of three different property vintages (old, new and modern building) across the UK regions, finding that the ripple effects are influenced by not only the spatial dimension but also the property vintage dimension.

Few studies have concentrated on addressing the underlying reasons of this ripple effect phenomenon. Meen (1999a) concluded that migration, home equity effects, spatial arbitrage and spatial patterns in the determinants of house prices are four possible explanations for the regional ripple-type pattern in Great Britain. Gray (2012) concluded that spatial spill over of house price growth is not only determined by commuting or migration alone; information flows and expectations are likely to reinforce inter-district transmission.

Although house price overspill research has mainly been investigated at a regional level, few studies have explored house price diffusion at small geographical levels. Gray (2012) was the first to focus on the LA district level. In his research, global measures of spatial autocorrelation and local indicators of spatial association (LISA) are used to track the house price diffusion. According to the ecological fallacy (Sedgwick, 2011), the statistical analysis of grouping based on different hierarchies will show different patterns. This revealed a fine-grained diffusion pattern. Until now, no research has focused on house price diffusion at a spatial granularity finer than LA level, across the UK, which also applies to England. There is a real gap in knowledge related to spatial connectivity and how housing wealth may be transferred between areas.

2.4.2 House price variation at micro geographic scales

At micro geographic scales, housing is immobile and location affects its value, hence housing neighbourhood analysis has long been a traditional concern of researchers (Orford, 2002; Boyle and Kiel, 2001; Li and Brown, 1980). The hedonic method is the most commonly used method to estimate the relationship between house price and its influential factors (Li and Brown, 1980). Hedonic price theory is derived from Lancaster's consumer theory and Rosen's theoretical model (Lancaster, 1966; Rosen, 1974). It is based on the hypothesis "that goods are valued for their utility-bearing attributes or characteristics" (Rosen, 1974: 34). In general, property prices can then be defined as a traditional linear function of a series of housing influence factors. This influential factor is normally divided into three categories: housing structure, location and neighbourhood characteristics. Housing is a heterogeneous good in terms of characteristics relating to the structure itself, such as type of house, property size, building age, room numbers and so on (Cheshire and Sheppard, 1995; De Nadai and Lepri, 2018; Sirmans et al., 2005). Floor area as a measure of property size, is treated as the most important structural characteristic influencing the house price (Morancho, 2003; Orford, 2010). What is more, determining the underlying location value in urban land and housing markets is complex, as house price values are affected by a variety of location and neighbourhood attributes (Richardson, 2013). These normally include the classic element of urban economic models, which is accessibility (Mok et al., 1995; Osland and Thorsen, 2008; Shen and Karimi, 2015). Other location determined characteristics, such as the character of neighbouring households, localized traffic effects and the quality of the micro environment and local public goods, such as schools and open space, also contribute to house cost heterogeneity (Gibbons and

Machin, 2003; Kane et al., 2006; Morancho, 2003; Szumilo et al., 2017).

The traditional hedonic model uses the ordinary least squares linear regression (OLS) model to identify the nature of relationships among variables. It assumes that the coefficients of the independent variables are uniform across the study area and that the error term is independently and identically distributed normally. But, house price is spatially auto-correlated in small areas in the real world. Therefore, the traditional hedonic model does not consider the autocorrelation among the regression variables. Several advanced methods have been proposed to incorporate spatial structural instability or spatial drift into models (Leung et al., 2000). Geographically weighted regression (GWR) is normally used when observation in close spatial proximity to one another are correlated. The GWR hedonic models have been developed which seeking to take into account spatial effects (Löchl and Axhausen, 2010; Lu et al., 2011). Furthermore, an extension of spatio-temporal hedonic models has been developed with additional spatiotemporal lag effects of previous sales in the vicinity of each housing sale, to help account for this (Fotheringham et al., 2015; Gelfand et al., 2004; Smith and Wu, 2009). Huang *et al* (2010) examined the applicability of traditional hedonic models, temporally weighted regression (TWR), geographically weighted regression (GWR), and geographically and temporally weighted regression (GTWR) models using the same house price data in Calgary, Canada. They show that the traditional hedonic model can only estimate 77.94% of house value variance and TWR model and GWR model improve the forecast, but the a GTWR model is the best. Fotheringham's (Fotheringham et al., 2015b) case study in London also shows that GTWR hedonic model is the best choice among the GWR hedonic models. The above versions of GWR methods assume that all predictors influence the response variable operating at the same spatial scale. To relax this assumption, multiscale geographically weighted regression (MGWR) is proposed to allow the predictors to influence the response variable at different spatial scales (Fotheringham et al., 2017). The above versions of GWR models are widely used in house price variation research (Fotheringham et al., 2015a, 2015a; Huang et al., 2010; Wu et al., 2014), but this GWR method family has a computational challenge when observation numbers exceed 10,000 (Li et al., 2019).

Compared with the above GWR models-based research method, a multilevel model (MLM) is a better statistical tool that includes the autocorrelation among the regression variables and works for a large number of observations. MLM can take any hierarchical (clustered) structures present in the data into account and has the ability to deal with more than one geographical location simultaneously (Hox, 2017; Jones, 1991a; Leyland and Groenewegen, 2003). The MLM allows individuals belonging to the same group to be more alike than a random sample. Moreover, within the groups in any given level, MLM allows relationships to vary around the overall relationship for all individuals across all the groups (Jones, 1991a). To produce more reliable estimates for groups with small sample sizes, MLM shrinks the estimates toward the overall average (Steele, 2008a). The multilevel variance components model is a MLM without explanatory variables. In exploring house price variations, it is a useful tool as it simultaneously deals with mean house price and house price variance at different geographical levels. Meanwhile, it decomposes the total housing variance across the available levels in the model, which is useful in quantifying the extent of spatial effects on house prices. Jones (1991b) firstly applied a multilevel model to understanding house price variation in Southampton. He applied a three-level model based on 918 house sales records and discovered that multilevel models demonstrate a considerable improvement over the traditional linear hedonic price modelling. Jones and Bullen's (1993) research further proves the advantages of using multilevel modelling, since it offers an improved description of the complexity of house price variation. This research was based on the 5 percent sample of mortgage completion data in the South East and the South West regions. Moreover, using a two-level model based on London's house price data, Jones and Bullen's multilevel model recognises that house price clusters within districts (Jones and Bullen, 1994). Since that work, there has been a continual growth in research using a multilevel model to explore house price variation across the world (Dong et al., 2015; Goodman and Thibodeau, 1998; Leishman, 2009). For example, Orford (2002) applied multilevel modelling to estimate the effects of location upon house prices in Cardiff, suggesting that the overall house price variation is composed of variations within districts, within communities and across individual properties. Recently, Feng and Jones (2016) were the first to present house price variation at five geographical scales in terms of postcode geography and census geography. These two geographical classifications show that house price, in terms of TP in London, has a hierarchical nature and that it is highly clustered at smaller geographical scales.

House prices for individual properties are frequently aggregated to larger spatial units, such as regions. For example, the Nationwide house price index regional quarterly series is created by the Nationwide Building Society (Nationwide Building Society, 2019). This is a theoretical average house price based on Nationwide mortgage data with consideration of a set of housing characteristics, in what is usually described as the mix-adjustment method (Nationwide Building Society, 2015). The official house price regional index, released by the Office for National Statistics (ONS), also uses a mix-adjusted approach to house prices, but uses LR PPD (Office for National Statistics et al., 2016). These two house price indices show the extent to which house prices differ between time periods at a regional level in England. However, they are not comparable as there are differences in data and the sets of housing characteristics considered. House price statistics at different geographical levels that use the same dataset may still be non-comparable due to different measurements. For example, the house price regional index and house price statistics for small areas are two house price series released by ONS (ONS, 2012b, 2018b). Both of these use LR PPD, but with two different methods: a mix-adjusted approach and a median approach. Lack of comparability can be an issue when studying house price movements and may be confusing to decision-makers.

To date there has been no systematic quantitative analysis to assess house price variation at a range of geographical scales across the whole of England. Some systematic quantitative analysis has been conducted to assess the TP variation in one city, such as in London (Feng and Jones, 2016; Law, 2017) and in Cardiff (Orford, 2002; Wang et al., 2015). In these studies, house price is normally presented as a TP. Furthermore, house price in the UK is normally aggregated at a given administrative geography level such as LA (ONS, 2015a) or MSOA level (ONS, 2017c), but local variations in stock composition and other factors mean that crude aggregation to geographic units for the purposes of studying price variations is problematic. Recently one solution has been to examine house price patterns by normalising the price per square metre. Powell-Smith (2017) was the first to map HPM across England at postcode level. Later in the same year, ONS launched an investigation into HPM at the LA level, but there has been no subsequent update of this (ONS, 2017d). These two investigations represent a valuable contribution to the knowledge of house price variation, but are single geographical level studies. Developing a comprehensive and systematic house price analysis at a variety of geographic scales should aid both government and public understanding of housing inequality and affordability issues in England (ONS, 2017a).

2.4.3 Section discussion

Modelling of English house price changes dates back to the 1970s (Ball, 1973; McAvinchey and Maclennan, 1982). The majority of housing research has explored the variation at coarse scales, such as regions or, conversely, in a specific city. The region-based research mainly focuses on house price changes within and between regions. The city-based research mainly focused on exploring the determinants of the spatial and temporal variation of property prices rather than the house price trends themselves. What is more, for housing research in England there is a lack of investigation of house price variation at different geographical and temporal scales, especially for the period after the global financial crisis (GFC) of 2008 (Cooper et al., 2013; Gray, 2012). Some recent studies have begun to address this (Feng, 2016; Gray, 2012; Law, 2018; Orford, 2017), but only a few have carried out this analysis nationally (Cooper et al., 2013; Feng, 2016; Gray, 2012). House prices vary not only across geographies but may also vary with a series of factors. There is insufficient understanding of the spatial extent of variations in housing markets in England (Holly et al., 2011). This lack of understanding means that important research questions about how the housing market functions across space and through time are difficult to answer, leading to differing views and intellectual traditions and uncertainty for policy makers. Furthermore, aggregate statistics for house prices at large geographical scales will mask variation at small geographic scales. Thus, simplistic measurements of housing affordability, based on house price and income at macro geographic scales, could hide the real picture of housing affordability within an area. This means that an in-depth analysis of house price variation at small geographical areas is extremely valuable to advance understanding of the spatial and temporal patterns of housing affordability and is likely to be useful as information for strategic policy decisions designed to improve the affordability of housing.

2.5 Research questions

Based on the above literature review, to achieve the research aim of this thesis, the research questions are defined as:

- To what extent does residential house price vary at small geographic levels in England, and how can we best characterise this variation?
- Could the analysis of house price variation at small geographic levels, combined with housing budgets for different types of buyer, help advance our understanding of the spatial and temporal patterns of housing affordability?

In order to address the research question, a range of data analysis and modelling options are conducted to support answering a series of substantive research questions. These are listed in Table 2.3.

Table 2.3 Chapters and corresponding research sub-questions

Chapter	Research sub-questions
Chapter 2: Literature review: house price and housing affordability	1. What are the current views on spatial house price variation across England? 2. How is housing affordability currently measured? 3. Does any literature exist on the relationship between housing affordability and house price at small geographic levels in the UK context? If it exists, what are its limitations? If not, what might be the reasons for this?

Chapter	Research sub-questions
Chapter 3: Measuring house price and housing affordability: a data review	1. What data are currently used to assess the spatial variation of house price in the owner occupied housing market? What are the limitations of these data particularly with regard to access and geographic scale? 2. What data are currently used to assess the temporal variation of house price? What are the limitations of these data particularly with regard to access, and time period availability? 3. What are the advantages and disadvantages of using UK official residential house price data (Land Registry Price Paid Data) compared with other house price datasets existing in the UK? 4. To what extent do problems with existing data (in terms of analysis at small geographical levels) relate to gaps in the availability of house price data? Could these problems be fixed by linking existing open datasets, or are other approaches likely to be required? 5. When linking open datasets, are any data lost? If so, how does this affect the data quality and potential results? What kind of test/method can help us to identify the information lost (e.g. Chi-square test and olmogorov–Smirnov test or J-divergence)? How can these tests identify which is the best period for observing for the research? 6. What datasets are currently being used to understand housing affordability in England and how is housing affordability measured?
Chapter 4: Understanding house price variation in England: a multi-scale exploration	1. How does residential house price vary at different geographic levels (i.e. LA, MSOA, LSOA) in England? 2. Given the above, at which level does most house price variation occur? Is this level an appropriate geographical scale to understand house price variation? 3. To better reflect house price variation, which available house price measure (e.g. transaction price, house price per square metre) should be used?
Chapter 5: Delineating the spatio-temporal pattern of house price variation by local authority in England: 2009 to 2016	1. With a focus on appropriate geographical scales (e.g. LA level), how does house price vary across various time scales? 2. What are the commonly used time scales for current house price statistics, and how does house price vary in these time scales? 3. Which is best time scale to explore the spatial and temporal patterns of house price variation to support the understanding of housing affordability? 4. What is the spatial temporal pattern of residential house price at LA scale?
Chapter 6: A new insight into local housing affordability in	1. How does house price change by property type at given geographical scales (e.g. LA level)? 2. Can a new housing affordability metric be created by

Chapter	**Research sub-questions**
England through further exploration of house price variation	considering house price variation at the most appropriate geographic level, and for housing budgets of different types of household? 3. To consider how scenarios can be used to simplify the variety of possible household compositions for the new housing affordability metric? 4. How does the housing affordability vary across different types of buyer and how does housing affordability change across space and time for a given type of buyer?
Chapter 7: Thesis discussion and final conclusions	1. Combining a knowledge of the spatial and temporal pattern of house price variation and housing affordability, what policy recommendations can be offered in order to address the current housing affordability issue? 2. What are limitations of the research and future research?

Chapter 3 Measuring house price and housing affordability in England: a data review

3.1 Introduction

House price data in England is imperfect (Gibb and Bailey, 2016; Wood, 2015) and this poses significant practical problems in exploring house price variation across England. Many readily available house price statistics are normally presented at a macro-geographic scale (i.e. region or LA), while house prices actually show spatially heterogeneous patterns at small geographical scales (ONS, 2016, 2017c). It is necessary to explore house price patterns at smaller geographic levels to gain a better understanding of the UK housing market. To support this, the choice of the dataset is regarded as critically important, but there has been little discussion of this in the literature (Gibb and Bailey, 2016; Whitehead et al., 2008; Wood, 2015). Meanwhile, the current official house price dataset (LR PPD) covers all residential transactions in England and Wales since 1995, and includes information on a number of housing characteristics, but it does not contain any accurate housing size information, such as floor area. House price data linked with information on individual property characteristics are difficult to obtain within UK (Gibbons and Machin, 2003; Orford, 2010), but dwelling size is regarded as one of the most important determinants of house price variation in house price modelling (Office for National Statistics et al., 2016; Orford, 2010). Building a comprehensive housing price database will produce an advanced understanding of house price variation.

Presently, there is no comprehensive database which contains TP along with property characteristics in England (Wood, 2015). This chapter reviews the available residential house price data and income data that are currently used in understanding the house price and housing affordability patterns in England. Section 2 provides an overview of house price datasets used in England with consideration of data content, time period, geographic coverage, geographical resolution and available data. The most comprehensive house price dataset will assist with answering the first research question: to what extent does residential house price vary at small geographic levels? LR PPD is chosen not only because it comprehensively records actual residential transactions but it is also more reliable to use at small geographic scales. A description of the LR PPD is shown in Section 3. Deficiencies in LR PPD are addressed and overcome in Section 4. Two data linkage methods are created to achieve data integration. Match rate, two statistic tests (Kolmogorov-Smirnov test) and one differences measure (J-divergence) are used to identify the transaction information that is lost after the data linkages. Based on the amount of lost information, this chapter identifies the most appropriate period from which to take input data for the research. After this, Section 5 outlines the available income data resources across England and its usefulness for estimating housing affordability. The chapter ends with the selection of the most appropriate income data for this thesis.

3.2 House price datasets in England

Table 3.1 Summary of current residential house price datasets in England

House price Dataset	Temporal coverage	Stage of recording transaction	Spatial coverage	Temporal coverage	Smallest geography of data	Data access	Used to calculate the index	Small area estimates
UK Residential Market Survey	1978-now	Monthly survey of 450-500 UK surveyors (residential sales and lettings)	UK	1978-now	Sample survey	RICS states that they support academic research and will supply the full data on a complementary basis (upon application)	None	Sample size too small
EGI (estates gazetts) database	1996-now	Latest asking price	London	1996-now	Postcode level	Open data	None	Sample size too small
Regulated Mortgage Survey	1969-now	Mortgage approval	UK	1969-now	Unknown	Only for Council of Mortgage Lenders members and associates	UK HPI (before 2010)	Sample size too small
Nationwide mortgage lending	1973-now	Mortgage approval	UK	1973-now	Postcode	Researchers need to apply	Nationwide Index	Sample size too small
Halifax mortgage lending	1983-now	Mortgage approval	UK	1983-now	Unknown	-	Halifax House Price Index	Sample size too small
Rightmove data	2001-now	Advertised asking prices	England and Wales	2001-now	Building level	-	Rightmove House Price Index	-
Zoopla data	2010-now	Advertised asking prices	England and Wales	2010-now	Address level	Open data through UBDC	None	-
LR PPD	1995-now	price paid for property	England and Wales	1995-now	Address level	Open data	UK HPI (after 010), LSL Acadata	-

There are eight main house price[3] resources within England (Marsden, 2015; Rae, 2015; Wood, 2015). Some of them are also used to construct house price indices by the UK government and some private organisations (Chandler and Disney, 2014; Jennings, 2018). Table 3.1 provides a summary of these house price datasets, along with house price indices that exist within England (Gibb and Bailey, 2016; Marsden, 2015; ONS, 2010, 2012b; Wood, 2015).

The UK Residential Market Survey conducted by the Royal Institution of Chartered Surveyors (RICS) is a monthly survey that investigated Chartered Surveyors' opinions on whether there was a change in house price over the previous three months (RICS, 2018). The sample size of the survey is quite small, normally lower than 500. Therefore, it is useful in providing a snapshot of national/regional housing market conditions and could give some anticipation of emerging market trends. It is not accurate presentation of the real extent of house price change and is unusable at small geographical levels. Meanwhile, the EGI (estates gazetts) database only contains London residential latest asking prices from 1996 and thus it is only useable to reflect the house price variance within London.

[3] All the house price in this chapter only covers the residential house price, but not rent price.

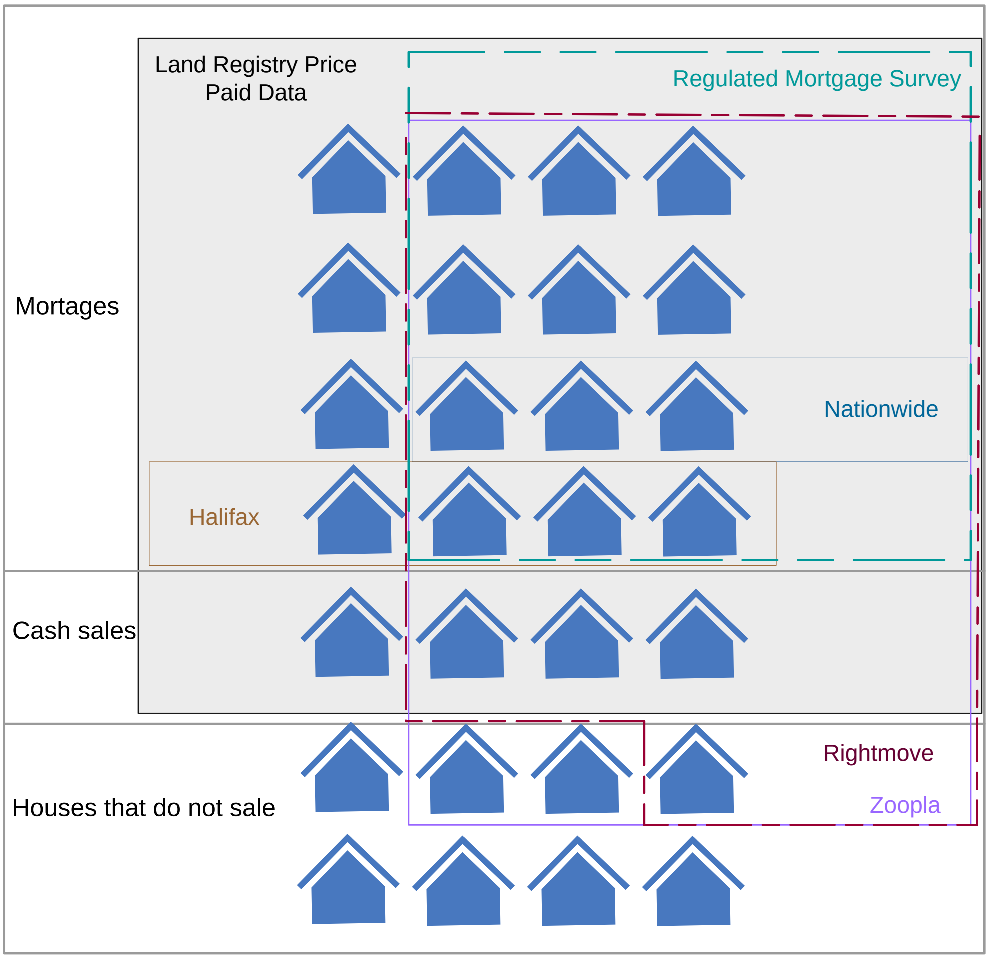


Figure 3.1 Data coverage of different house price sources

The other six house price datasets all cover the whole of England. Figure 3.1 demonstrates the data coverage of these six house price datasets. The Regulated Mortgage Survey by the Council of Mortgage Lenders, Nationwide and Halifax datasets covers property transactions that were bought through a mortgage. The Regulated Mortgage Survey is the biggest mortgage house price dataset in the UK as well as containing the longest time period. It covers all mortgage data from all UK mortgage lenders, occupying 75%-80% of the mortgage market submitted data (ONS, 2013). This dataset offers quite credible results to depict the mortgage housing market in England, but it is not publicly available. The other two mortgage datasets (Halifax mortgage lending and Nationwide mortgage lending data) are subsets of the Regulated

Mortgage Survey dataset. The Nationwide mortgage lending dataset is smaller than the Halifax mortgage lending dataset (Jennings, 2018), but it is open for academic research. Therefore, there is a large amount of research that uses the Nationwide mortgage lending dataset to explore the house price dynamics within the UK (Ahlfeldt et al., 2012; Law, 2017; Lu et al., 2014).

In the real world, some people purchase dwellings with cash only, as opposed to using a mortgage. Using the mortgage lending datasets to conduct research or create house price statistics may be biased if the sample does not include similar houses to those purchased with cash. Only when the houses purchased with cash behave similarly to those purchased with mortgages and this pattern remains the same over time, is it acceptable to use the mortgage house price dataset to detect the real house price variance of the housing market. However, these conditions are unlikely to hold and unable to be controlled for. The LR PPD can directly overcome this shortcoming as it covers the both mortgage transaction and cash transaction. In addition, it is an open data resource and records transactions at address granular. It could credibly be used to analyse patterns of residential housing market at any given geographical level across England.

Alternative datasets such as Rightmove and Zoopla also offers datasets, they use advertised asking house prices and have a large real time data sample. Advertised house price can be a poor indicator of house price as it may be different to the final TP. Moreover, similarly as with mortgage house price data, it may show misleading patterns when detecting the house price variance over time. Properties advertised on the Zoopla and Rightmove websites may not result in successful sales. Thus, advertised asking house price from Rightmove and Zoopla shows less reliable house price data when comparing to LR PPD.

LR PPD shows a significant advantage in data coverage because it comprehensively records the actual residential transactions. This also means it will be the most reliable dataset to conduct house price statistics at small geographic units, but the LR PPD release data with a quarterly registration lag (ONS, 2018b). This registration lag may cause issues when trying to identify house price trends in the most recent quarter. Thus, LR PPD is a relatively reliable dataset to reflect the history of house prices.

3.3 LR PPD

LR PPD is an administrative dataset from the Her Majesty's LR, which became open access in 2013 (HM Land Registry, 2015). This records almost all the actual residential transactions since 1995 at address level with several sale types excluded. Although the LR PPD omits some types of residential property sales (e.g. sale through the government's 'right-to-buy' scheme), it still provides the most accurate picture of residential property sales at full market value in England and Wales (HM Land Registry, 2016; Marsden, 2015; South and Henretty, 2017). The ONS uses this data to calculate certain house price statistics, such as House Price Statistics for Small Areas (South and Henretty, 2017) and the Official House Price Index (Office for National Statistics et al., 2016). Table 3.2 shows an explanation of data items in the LR PPD. The dataset not only contains the property sales price, transaction date and property address information, but also shows house type (detached, semi-detached, terraced houses and flats/maisonettes) and tenure (freehold/leasehold), and whether a property is newly built or whether it was sold at full market value.

Table 3.2 Explanations of information fields in LR PPD[4]

Data Item	Explanation
Transaction unique identifier	A reference unique number which is recording each published sale. e.g. {955B1020-9223-4981-AFF1-72C47E6CC60E}
Price	Sale price (transfer deed). e.g.10,000
Date of transfer	Date when the sale was completed. e.g. 2006-10-13
Property type	Indicates the type of house: D = Detached, S = Semi-Detached, T = Terraced, F = Flats/Maisonettes, O = Other
Old/New	Indicates the age of the property and applies to all price paid transactions, residential and non-residential. There are two categories: a newly built property, an established residential building. If the property is firstly sold since 1995 it will identify as ‘a newly built property’. Y = a newly built property, N = an established residential building
Duration	The tenure of property: freehold, leasehold
PPD category type	Indicates the type of Price Paid transaction. A = Standard Price Paid entry, includes single residential property sold for full market value. B = Additional Price Paid entry including transfers under a power of sale/repossessions, buy-to-lets (where they can be identified by a Mortgage) and transfers to non-private individuals. Category B is identified from October 2013.
Postcode	e.g. WC1H 9QH
PAON	Primary Addressable Object Name. such as the house number or name. e.g. 36
SAON	Secondary Addressable Object Name. Where a property has been divided into separate units (for example, flats), the PAON (above) will identify the building and a SAON will be specified that identifies the separate unit/flat. e.g. Flat 302
Street	e.g. Tottenham Street
Locality	e.g. London
Town or city	e.g. London
District	e.g. Camden
County	e.g. Greater London
Record status	Indicates additions, changes and deletions to the records A = Addition; C = Change; D = Delete. e.g. A

As shown in Table 3.2, the LR PPD offers property sales prices and transaction date information. This is quite useful in exploring the residential house price, in a given period based on the transaction date. The LR PPD used in this research was downloaded on

[4] Resource: https://www.gov.uk/guidance/about-the-price-paid-data#data-excluded-from-the-house-price-index-and-price-paid-data

14/9/2017, and records 22,578,068 transactions in England and Wales between 1/1/1995 and 31/7/2017.

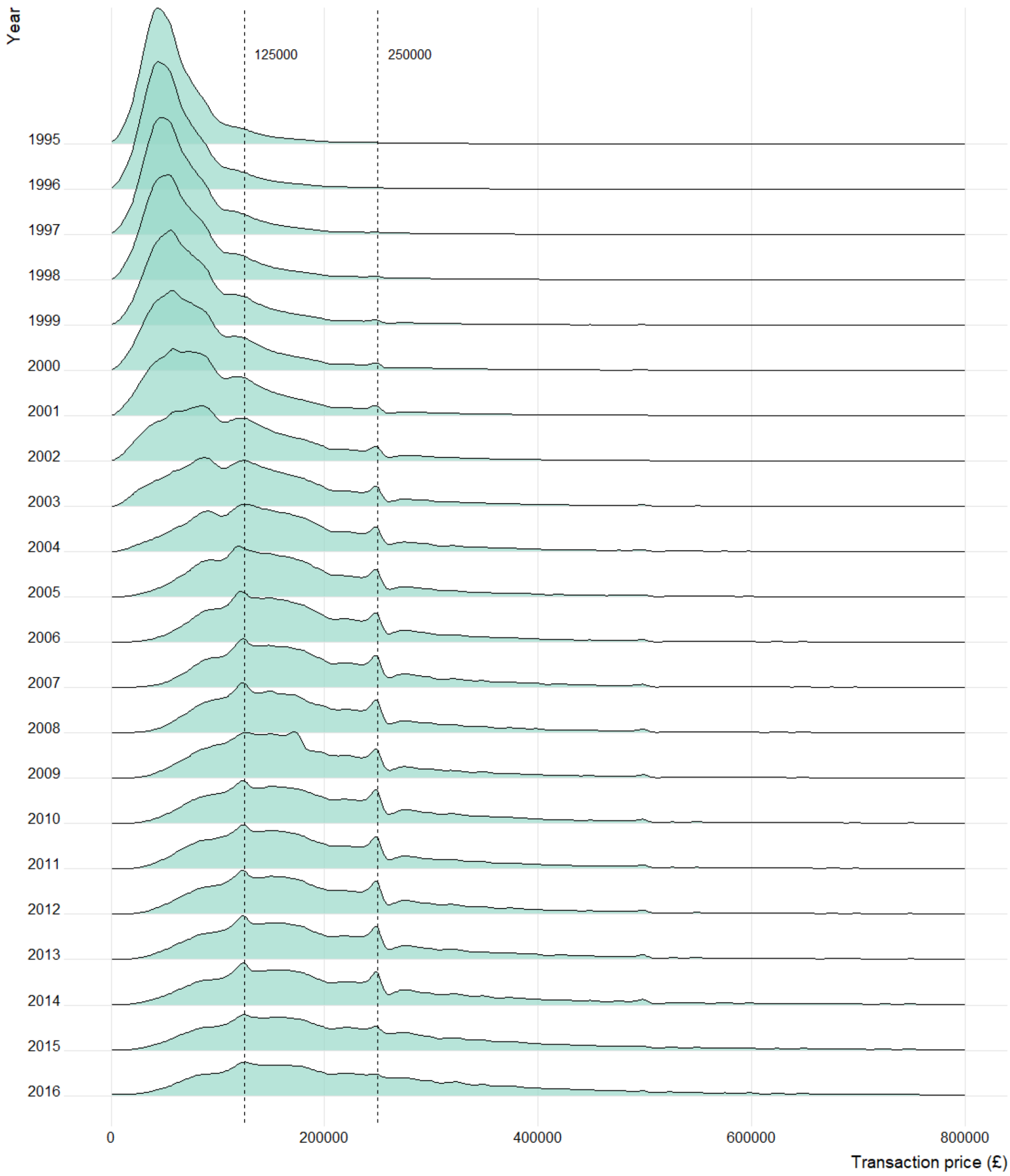


Figure 3.2 A Joyplot version of TP density plots in England and Wales,1995-2016[5]

Figure 3.2 shows the TP distribution from 1995 to 2016. Over this period, TP

---

[5] The LR PPD covers the period from 1/1/1995 to 31/7/2017. It does not cover the whole transactions occur in 2017. Thus all the description analysis within this section below not include the transactions in 2017. As the house price distribution shows a long tail and this figure only plots the below £800,000 part.

distributions in each year are seen to be positively skewed. It means prices are mainly clustered around a relatively low value together with a few extreme high values. Meanwhile, TPs have become increasingly dispersed over time as the overall range of TP has widened dramatically during the last 22 years. The two local peaks (at £125,000 and £250,000), that may be observed in the graphs since 1998, reflect the Stamp Duty Land Tax (SDLT) thresholds. Moreover, TPs after 2006 exhibit a new peak at £500,000, which is also SDLT related.

The average number of annual transactions in England and Wales from 1995 to 2016 is around one million. Figure 3.3 shows how the transaction volume has changed from 1995 to 2016. There is a significant turning point when the GFC erupted in 2007. Transaction numbers show a generally increasing trend from 1995 to 2007, but this suddenly decreases by about a half in 2008. The number of residential property sales continues to recover after 2009, with an increase to over one million after 2015.

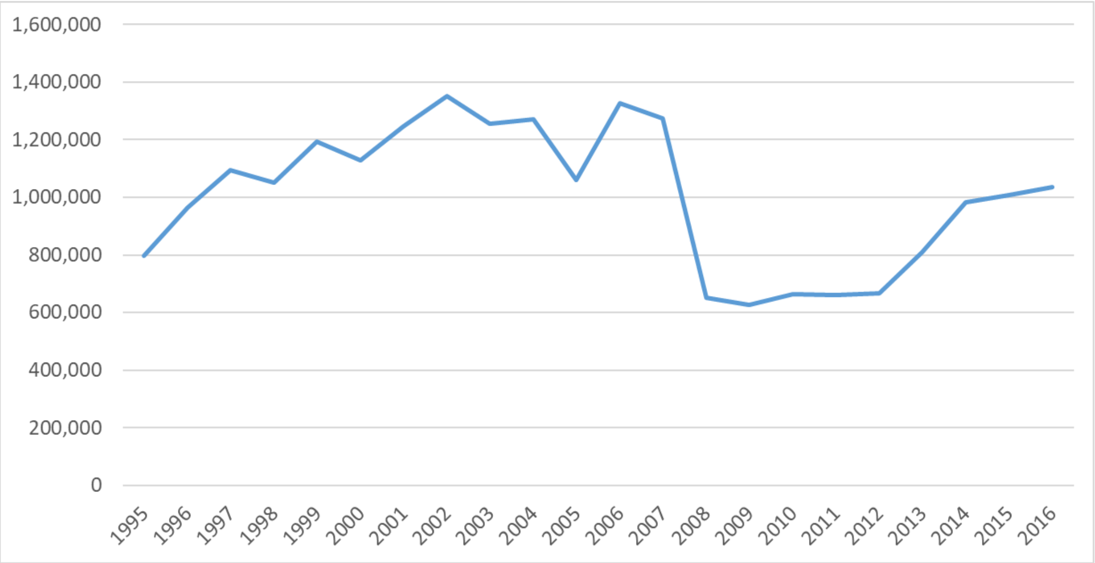


Figure 3.3 Transactions sales change in England and Wales, 1995-2016

LR PPD also records the property’s postcode information. This can be linked with the

National Statistics Postcode Lookup (NSPL)[6] to directly add in a series of statistical geographies, such as region, administrative, electoral and health statistical geographies. Figure 3.4 presents the transaction sales distribution among regions in England and Wales for the whole time period from 1995 to 2016. It is obvious that the transactions in Wales are lower than in England. Welsh transactions are always the smallest during the 22 years. Furthermore, the trend of transactions in Wales shows similarity to the trend in the North East of England. Looking at the transactions in England, the North East always has the lowest transactions in England from 1995 to 2016, while the South East always has the highest transactions in England in the same period. Besides, when looking at each region separately in England between 1995 and 2016, each region generally shows a similar transaction trend.

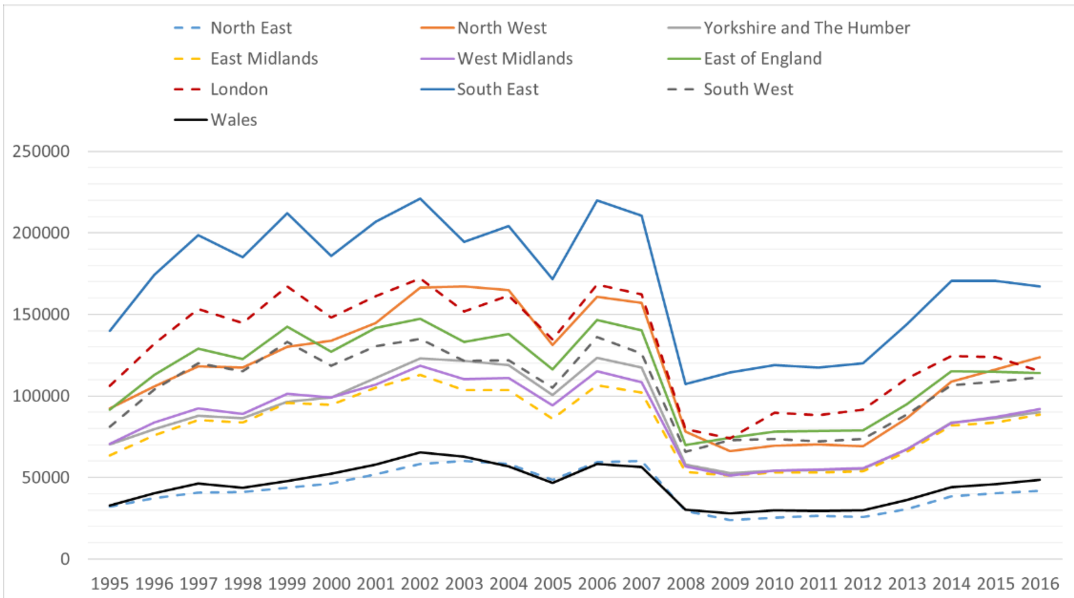


Figure 3.4 Property transactions sales by regions, 1995-2016

[6] National Statistics Postcode Lookup is a list of both current and terminated postcode, which contains all the postcode in the UK along with a selection of the statistical geographies they are situated within, it produced by ONS Geography to support the production of area based statistics from postcode data. Detail see: http://geoportal.statistics.gov.uk/datasets?q=National%20Statistics%20Postcode%20Lookup%20November%202017&sort=name

The transaction trend in England and Wales can be divided into four periods: the time period between 1995 and 2007, the time period between 2007 and 2008, the time period between 2008 to 2012, and the time period between 2012 to 2016. For the first time period (1995 - 2007) the transaction sales show a generally upwards trend with a small fluctuation down in 2005. As for the second time period (2007-2008), transaction sales show a large decrease, of almost 50%. In the third time period (2009-2010), transaction sales of all regions (except London) show a continuing similar low transaction sale trend while London shows an increasing trend. In the fourth time period, regions' transaction sales show a clear increasing trend from 2012 to 2014, this increase slows down during 2014 to 2016. Meanwhile, London's transactions drop a bit during 2015 to 2016.

The LR PPD not only contains the property sales price, transaction date and address information, it also offers house type (detached, semi-detached, terraced houses and flats/maisonettes) and property duration (freehold/leasehold), whether a newly built property and sold through full housing market value. This is useful in understanding house price variation according to the age of dwelling (old or new houses), different property types and duration (freehold/leasehold).

3.4 Enriching the LR PPD

Dwellings have heterogeneous characteristics and therefore the house price will differ, even within the same neighbourhood. Moreover, house prices show spatial sensitivity (Halket et al., 2015; Palm, 1978), meaning they varies across locations. That is why house price is normally presented at a certain location. House prices in the same neighbourhood tend to be similar to each other, but house prices vary as a result of physical attributes, such as dwelling size, age, structural design and historic value (Ahlfeldt et al., 2012; Goodman and Thibodeau, 1995; Kain and Quigley, 1970). Given this, the LR PPD has two potential limitations as a tool for understanding house price variation. One is that the data are not geocoded, the other is that they do not include the property characteristics (e.g. property size) information. Two methods are outlined below to overcome these two limitations. One method aims to geo-reference LR PPD at the building level, whilst the other aims to further add in dwelling characteristics (i.e. total floor area and number of habitable rooms) to the geo-referenced house price data. With the combination of these two methods, an enhanced house price database is created. To assemble the new database, three other datasets are used: OS MasterMap Topography Layer (MMTL); OS ABP; and Domestic Energy Performance Certificates (EPCs).

OS MMTL is an OS spatial dataset which represents individual buildings as geolocated polygons along with a unique geocode (TOID, Topographical Identifier) in the UK. OS ABP contains address information for current properties in the UK. For each current active property, OS ABP records the property's Unique Property Reference Number (UPRN), generated by the LA together with the OS TOID and Royal Mail postal delivery address. Linking on a property's geocode (TOID) from OS MMTL, to the same code in OS ABP enables the matching of a building's postal delivery address to its geographic information (i.e. coordinates). This facilitates the geo-referencing of the LR PPD at building level through transfer the transaction's home delivery address to the building's geographic information.

Domestic EPCs are held in an open administrative dataset in the Ministry for Housing, Communities and Local Government (MHCLG). The dataset records a property's theoretical energy performance, property address and its physical characteristics information, such as its total floor area and number of habitable rooms. Since 2008, EPCs are legally required in the UK when a building or building unit is offered for sale or rent and the certificates remain valid for 10 years. Property data in LR PPD and Domestic EPCs are able to be linked together based on the address information in each dataset.

Figure 3.5 shows the workflow of enhancing the LR PPD with the above three datasets, using two methods. Method 1 geotagged the LR PPD at building level. Method 2 enriches the spatial data with property size information from Domestic EPCs.

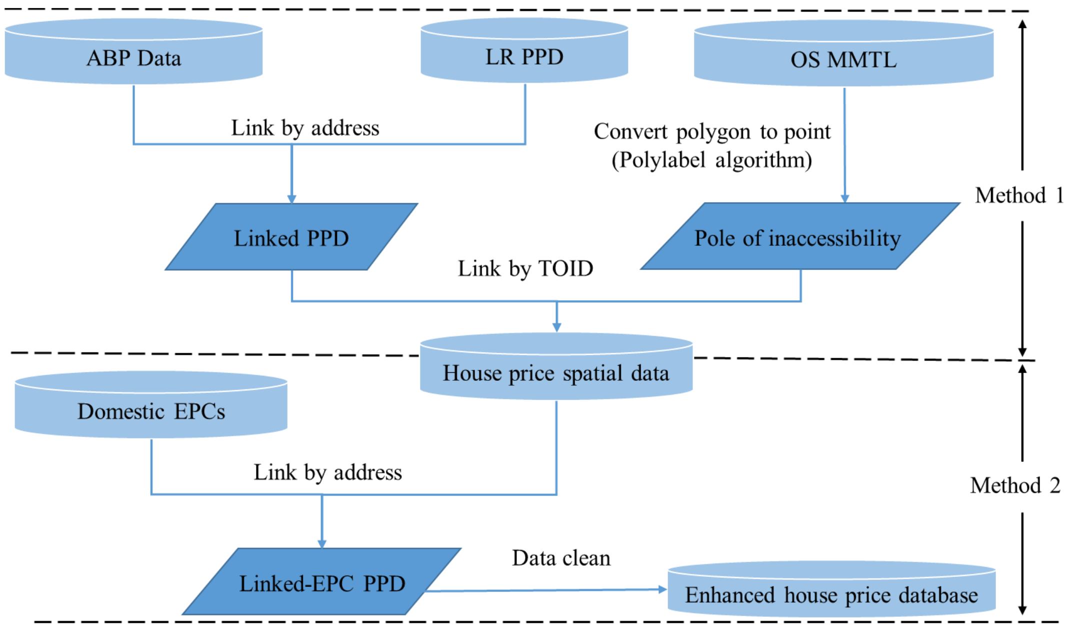


Figure 3.5 A brief flowchart of enhancing the LR PPD

3.4.1 Geotagging the LR PPD at building level

Geographic information exists in the form of the address string in the LR PPD. The NSPL is frequently used to link geographic information (i.e. latitude and longitude) to the LR PPD through matching the postcode (South and Henretty, 2017). This method cannot accurately pinpoint the dwelling's real location, since it only locates the postcode's centroid point. OS ABP contains a property's building polygon TOID in OS datasets and its dwellings' postal delivery addresses. OS MMTL contains the current building's polygons with TOID. Linking these two datasets through TOID creates a database that is interchangeable between the building's address and its geographic information. Therefore, geocoding the LR PPD can be achieved at the building location by integrating LR PPD with ABP and OS MMTL data.

The combination of LR PPD, OS ABP data and OS MMTL build a foundation for the geo-referencing of LR PPD process (method 1 in Figure 3.5). LR PPD and ABP data are first linked by address information (postcode along with address strings), then link back to the OS MMTL matching through the TOID. On the other side, an iterative grid algorithm called Polylabel (Garcia-Castellanos and Lombardo, 2007; Hügel, 2017) is used to calculate the pole of inaccessibility[7] of each polygon as a proxy of geolocation of the building. The last step is to link these three datasets using the TOID to build a house price spatial database.

[7] Pole of inaccessibility is a geographical point that represents the most remote place reached in a given area. The definition of pole of inaccessibility is the point within a polygon that is farthest from an edge. In cartographic visualization, it is used to position the text label on the centre of polygon.

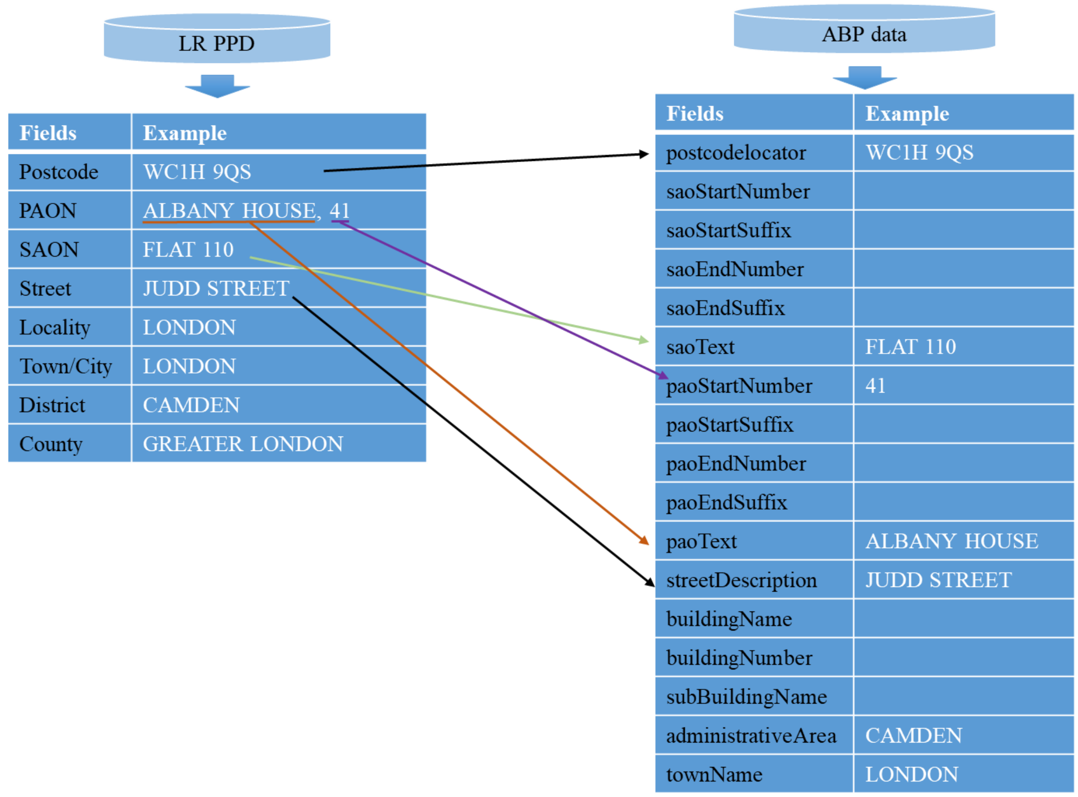


Figure 3.6 Address components difference in LR PPD and ABP data

Linking LR PPD with ABP by address information presents difficulties as the address records between these two datasets are structured differently (Figure 3.6). The full postal delivery addresses in the LR PPD are categorized into four address information items (i.e., postcode, paon, saon and street). The ABP data not only contains the same postcode and street records, but also includes building name, building number and sub-building name. Moreover, it divides PAO (Primary Addressable Object) information as ‘paostartnumber’, ‘paostartsuffix’, ‘paoendnumber’, ‘paoendsuffix’ and ‘paotext’. Similarly, SAO (Secondary Addressable Object) information divides in the same way, named as ‘saostartnumber’, ‘saostartsuffix’, ‘saoendnumber’, ‘saoendsuffix’, ‘saotext’ respectively. These differences mean that matching is not straightforward and a multi-stage process is required to achieve successively more matches.

Basic data cleaning and standardization are implemented to support the address-based data linkage. As shown in Table 3.3, thirty-two new address variables are created in either LR PPD or ABP data to support the data linkage, nine of these are created in the LR PPD and the rest of twenty-three new variables are created in the ABP data.

Table 3.3 New address variables created from existing address field[8]

Type	New variable name	Create method
Combine	SAONPAON	Combine SAON and PAON with a blank space
	PAONSTREET	Combine PAON and street with a blank space
	SAONSTREET	Combine PAON and street with a blank space
	bb	Combine buildingname and buildingnumber, using a comma
	pp	Combine paostartnumber and paostartsunffix
	pp1	Combing paotext and paostartnumber fields using a comma
	pp2	Combing paotext and pp fields using a comma
	pp4	Combine paostartnumber and paostartsunffix using hyphens
	ppp	Combine paotext and pp4 with a blank space
	ss	Combine saostartnumber and saostartsuffix
	ss1	Combine saostartsuffix and saostartnumber
	subss	Combine subbuildingname and ss with a blank space
	saopp	Combine saotext and pp with a comma and a blank space
	sp	Combine ss and paotext fields using a blank space
	ssp	Combine saotext and sp1 with a comma and a blank space
	saobui	Combine fields saotext and buildingname using a blank space
	psao	Combine the paostartnumber and saotext1
Stripping	PAON1	Stripping surrounding whitespace from hyphens and the comma in PAON field.
	PAON2	Stripping surrounding whitespace from hyphens in PAON field
	SAON2	Stripping surrounding whitespace in SAON field
	saotext1	Deleting the 'FLAT ' leading string in saotext
Prepend string	FLATSAON	Prepend the SAON with 'FLAT ' string
	FLATPAON	Prepend the PAON with 'FLAT ' string
	UNITPAON	Prepend the PAON with 'UNIT ' string
	flatsao	Prepend the saostartnumber with 'FLAT ' string
	flatss	Prepend the ss with 'FLAT ' string
	flatsub	Prepend the subbuildingname with 'FLAT ' string
	unitss	Prepend the ss with 'UNIT ' string

[8] Variables written as capitals are new variables added to LR PPD, the lower-case variables are new variables added to ABP data.

Type	New variable name	Create method
	flatpao	Prepend the paostartsuffix with ‘FLAT ’ string
	paostartnumber1	Prepend the paostartnumber with ‘FLAT ’ string
Replace	subbuildingnamenew	Replace ‘UNIT’ and ‘APARTMENT’ string in subbuildingname to ‘FLAT ’ string
	saotext2	Replace the ‘APARTMENT’ ,‘SUITE’ string in saotext to ‘FLAT ’ string and delete ‘.’ string in saotext

The data linkage between LR PPD and ABP data is designed to match within each unique postcode unit belonging to LR PPD. Some postcodes included in the PPD are not covered by the ABP data. The transactions with these postcodes are deleted first. A data linkage is created using a thirteen-stage process that has 97 matching rules; it is based on the address string fields shown in Figure 3.6 and Table 3.3. Details of the 13-stage process and matching rules are shown in Figure A1 and Table A1 (Appendix A1). Table 3.4 summarizes the match rate for each stage in method 1.

Table 3.4 Match rate for different stages

Stage	Match rate	Cumulative match rate
Stage 1	0.002%	0.002%
Stage 2	91.51%	91.51%
Stage 3	2.19%	93.70%
Stage 4	0.23%	93.93%
Stage 5	0.74%	94.67%
Stage 6	0.11%	94.79%
Stage 7	0.30%	95.09%
Stage 8	1.83%	96.91%
Stage 9	0.32%	97.24%
Stage 10	0.47%	97.70%
Stage 11	0.01%	97.71%
Stage 12	0.19%	97.90%
Stage 13	0.04%	97.94%

LR PPD data used here cover transactions between 1/1/1995 and 31/7/2017 in England and Wales. Using the 13 stage/97 rules model, 97.94 % of transactions (22,113,003) are successfully matched. This data linkage result is designated as the data link table as shown in Figure A1 and Figure 3.5. Stages 2 and 3 together achieving a 93.70%

match rate, without additional stages being performed. These two stages therefore constitute the main matching process. Given the differences in address string format between the LR PPD and ABP datasets, a more complete data linkage was achieved by processing the newly created address variables through another eleven stages termed the match cleaning up process.

Following the workflow in Figure 3.5, the data link table obtained from the 13-stage matching linkage contains a unique transaction identifier (*transactionid*) from the LR PPD and TOID (*ostopotoid*) from OS ABP data. Then using the LR PPD with the data link table we can successfully add TOID to the TP to give the linked PPD. After that, the linked PPD can be geo-referenced by linking the building's centre point (Pole of inaccessibility) by TOID. The method 1 process (Figure 3.5) successfully geo-referenced 22,019,341 records at building level and this new dataset is designated as the house price spatial data.

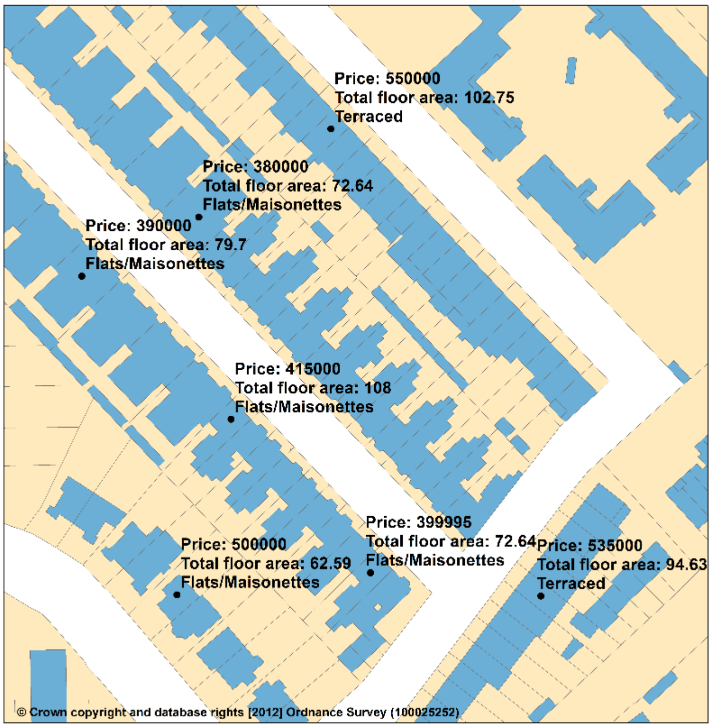


Figure 3.7 Sample of house price spatial data with OS Master Map

A sample of house price spatial data is shown in Figure 3.7. Each black solid point represents one record in house price spatial data. There are two major advantages in this spatial dataset. Firstly, unlike the original LR PPD data, house price spatial data can now be aggregated at the level of any geographical unit (e.g. street level). Secondly, fully georeferenced house price data is more analytically flexible than data represented at postcode unit by linking to the NSPL. This flexibility allows for a much wider range of spatial analyses to be conducted, such as exploratory spatial data analysis and spatial interpolation.

A 100% match rate is not to be expected mainly because in both datasets the addresses are structured differently. Additionally, there are three other reasons. Firstly, 0.12% of the LR PPD lack the postcode information in the price paid dataset. Secondly, some transactions do not possess matching address information in the ABP dataset; this may be because these properties no longer exist. Thirdly, some transaction address records are insufficiently detailed to identify the unique TOID in which they are situated. This issue caused one-to-many relationship problems with one (transaction) being related to many buildings during the matching process.

3.4.2 Enrichment house price data spatial data with property size information

Modelling suggests floor area is the most important determinant of house price (De Nadai and Lepri, 2018; Morancho, 2003; Orford, 2010; Sirmans et al., 2006; Thwaites and Wood, 2005). Thus, enriching LR PPD with floor area information will be highly valuable in supporting house price analysis, especially for house price variation analysis. Some researchers have started to use the combination of LR PPD and EPC data to undertake house price research. The first HPM map in England and Wales is created based on TP from LR PPD and each property's total floor area from EPCs (Powell-Smith, 2017). This is achieved by linking the LR PPD since 2007 with EPCs in England and Wales. This map offered a new insight into HPM patterns at postcode district level with linked LR PPD. Moreover, Fuerst et al (2013) combined LR PPD and EPC data to explore the relationship between energy performance and house prices across the UK in the period from 1995 to 2011. These two researches show an achievable approach to enrich LR PPD with variables in EPC data, which can be used to enrich the house price spatial data with the total floor area information from Domestic EPCs. Although these two researches successfully linked the LR PPD and Domestic EPCs, the detailed method is inaccessible. Consequently, this research created its own address-based method to link between house price spatial data and Domestic EPCs. This section has described an address-based method (method 2 in Figure 3.5), which aims to enrich the house price spatial data with the total floor area information from Domestic EPCs.

3.4.2.1 Data linkage

The EPC dataset used in this study is the first version downloaded on 31/5/2018. The first version EPC contains 85 items with 15,623,536 Domestic EPCs from 1/1/2008 to 1/10/2016. It is the only available version before MHCLG released the second version in 2019. After 2019, MHCLG started to release updated versions two or four times per year. However, the updated versions no longer contain records for EPCs lodged prior to 1/10/2008. Given this constraint, this thesis chooses use the first version (1/1/2008 to 1/10/2016). Table 3.5 shows the description of the key property characteristics recorded in Domestic EPCs.

Table 3.5 Explanations of address string and key property characteristics in EPC data[9]

Item	Explanation
Address1	First line of the address. e.g. Flat 110.
Address2	Second line of the address. e.g. Albany House.
Address3	Third line of the address. e.g. 41 Judd Street.
Postcode	The postcode of the property. e.g. W1T 4RW.
Property type	Describes the type of property. e.g. Maisonette, Flat, House, Bungalow, Park home.
Built form	The building type of the Property e.g. Enclosed End-Terrace, Detached, End-Terrace, Semi-Detached, Mid-Terrace, Enclosed Mid-Terrace.
Inspection date	The date that the inspection was actually carried out by the energy assessor.
Lodgement date	Date lodged on the Energy Performance of Buildings Register.
Total floor area	The total useful floor area is the total of all enclosed spaces measured to the internal face of the external walls, the gross floor area as measured in accordance with the guidance issued from time to time by the Royal Institute of Chartered Surveyors or by a body replacing that institution.
Floor level	Flats and maisonettes only. Floor level relative to the lowest level of the property (0 for ground floor). If there is a basement, the basement is level 0 and the other floors are from 1 upwards.
Number of habitable rooms	Habitable rooms include any living room, sitting room, dining room, bedroom, study and similar; and also a non-separated conservatory. A kitchen/diner having a discrete seating area (with space for a table and four chairs) also counts as a habitable room. A non-separated conservatory adds to the habitable room count if it has an internal quality door between it and the dwelling. Excluded from the room count are any room used solely as a kitchen, utility room, bathroom, cloakroom, en-suite accommodation and similar; any hallway, stairs or landing; and also any room not having a window.
Floor height	Average height of the storey.
Address	Field containing the concatenation of address1, address2 and address3.

Figure 3.8 represents the process of data linkage between house price spatial data and Domestic EPCs. These two datasets offer the property information at address level, but their address structures are different. Basic data standardization is conducted

[9] Resources: https://epc.opendatacommunities.org/docs/guidance

before linking house price spatial data and Domestic EPCs. First, all the address strings in the Domestic EPCs were capitalised and then new address variables were created separately in the house price spatial data and Domestic EPC data sets. Finally, the newly created address variables were used to achieve the data linkage. Following this process, 180 new variables were created in the house price spatial data and 95 new variables were created in the EPC data to assist the data linkage. Details of the new variable creation methods are shown in Table A2 (Appendix A2).

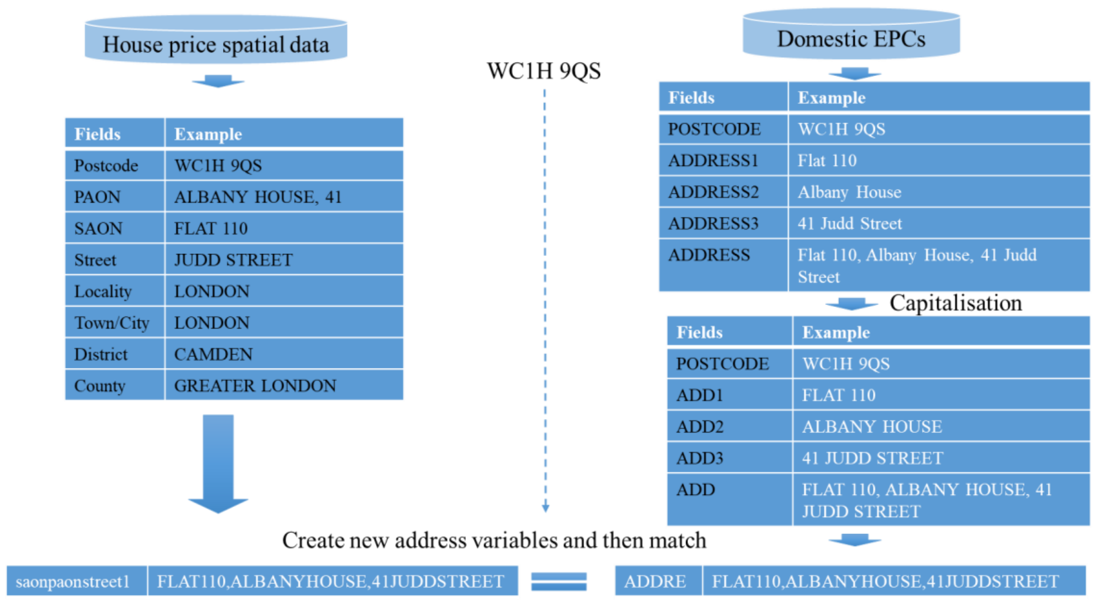


Figure 3.8 An example of data linkage process

Before the matching, transactions without postcodes in the Domestic EPCs dataset were excluded. A total of 0.64% of the data was deleted after applying this rule. Then, with the newly created address variable in Table A2, a matching method containing a 4-stage (163 matching rules) matching process was designed to combine the house price spatial data and Domestic EPCs. Details of the matching process and matching rules are shown in Appendix A3. Following the combination of house price spatial data and Domestic EPCs, 14,519,565 geo-referenced transaction records were successfully linked with an EPC.

Within the linked EPC data, 13,881,493 of the entries are transactions in England. The match rate of transactions in England is shown in Figure 3.9. The matching rate between 2009 and 2016 is higher than 90%, while the matching rate of the rest of the period is lower than 70%. As the first version of the EPC data only covers the period between 1/1/2008 and 1/10/2016, the match rate is relatively high (over 90%) for the same year period (2008-2016). After checking the transactions (2008-2016) which failed to link, it was found that there are some sold dwellings which were not recorded in the publicly available EPC data. This makes 100% matching unachievable. The matching rate of the period before 2008 and after 2016 is in the range of 50% to 70%. This is mainly due to the dwellings sold before 2008 or after 2017 having also been sold again or rented during 2008 to 2016, permitting them to be matched in the Domestic EPC.

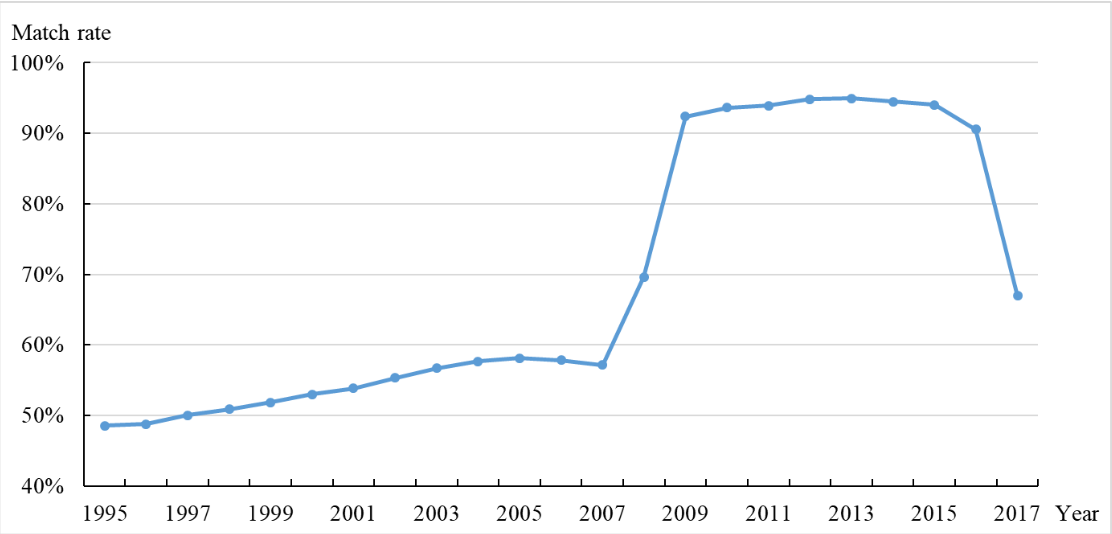
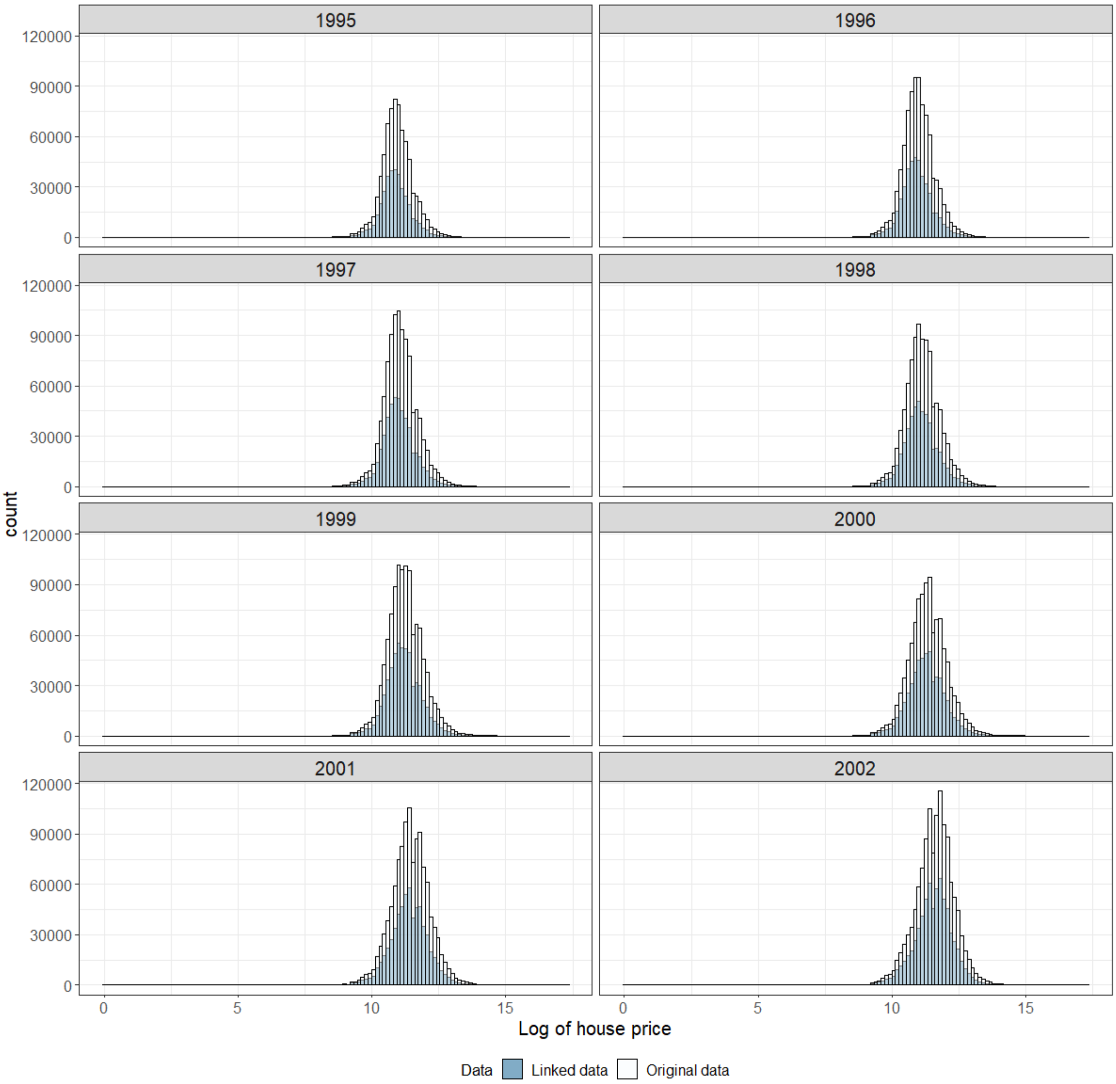


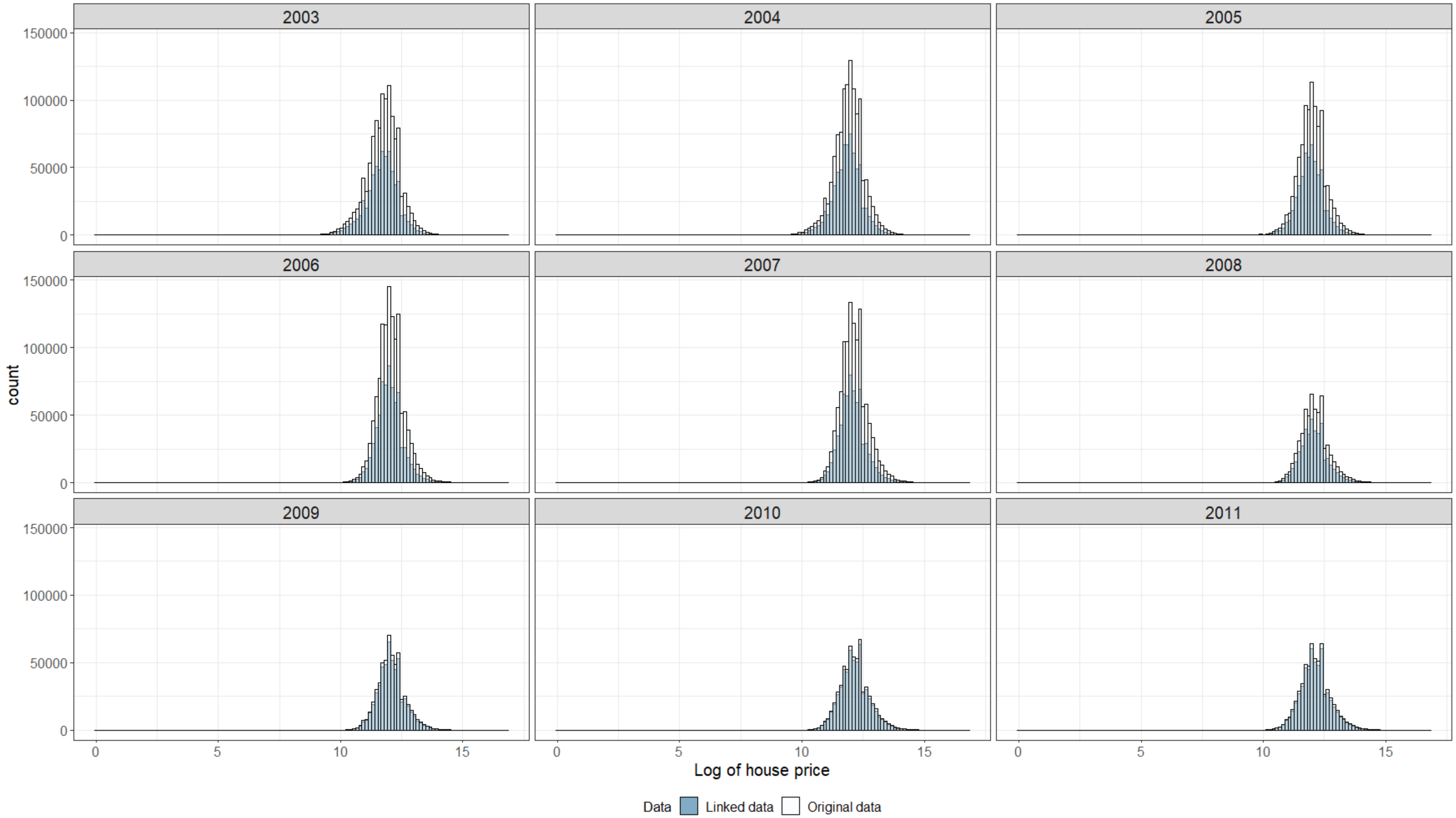
Figure 3.9 Match rate of LR PPD in England,1995-2017

3.4.2.2 Evaluation of house price information lost after data linkage

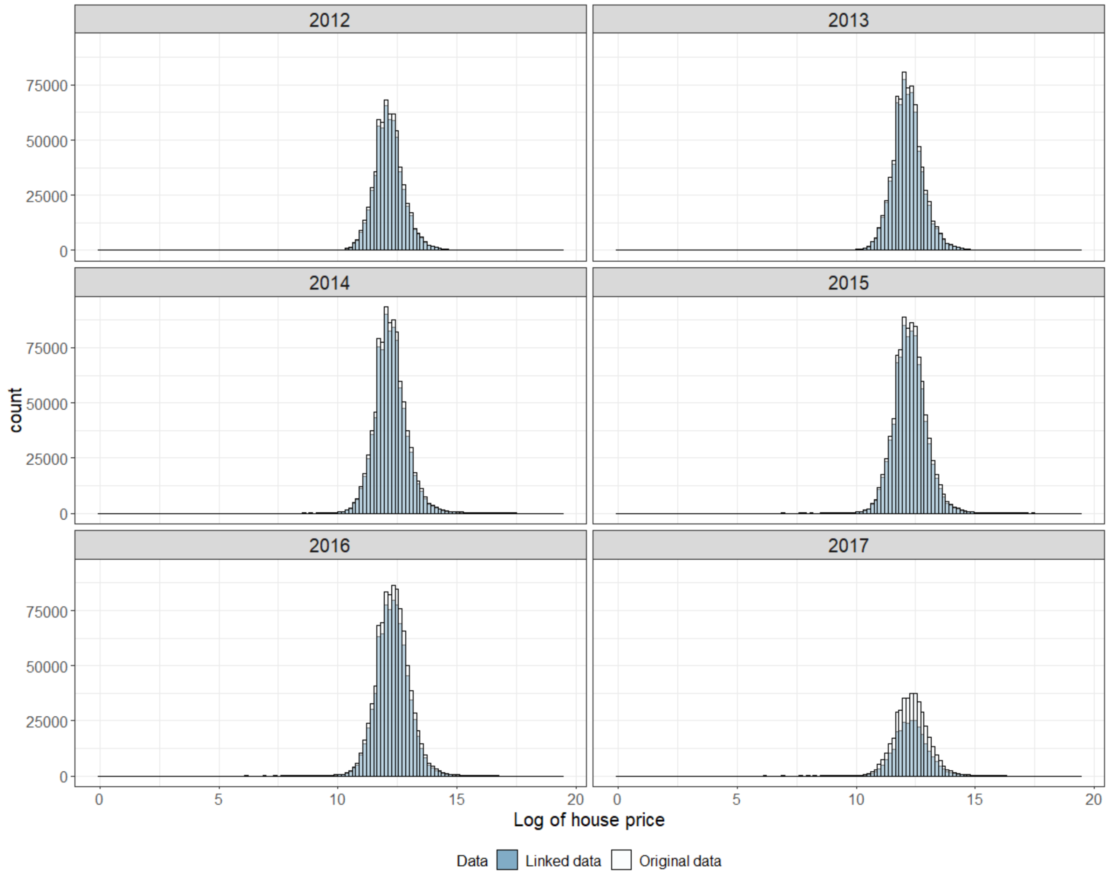
Match rates offer a crude way to quantify the matching performance, but visualization of the house price difference before and after linkage displays a clear picture of the matching performance by considering all the available house price values in the dataset. As the house price distribution follows a positive skew distribution with a long tail (Figure 3.2), the logarithm of house price is used to rescale the house price range. Histograms of the logarithm of house price from the transaction data in house price spatial data (geo-referenced PPD) and linked-EPC PPD in a certain given year is chosen to visualise the house price distribution change (Figure 3.10). In each graph, the distribution of the linked-EPC PPD (linked data) is overlaid onto the distribution of the house price spatial data (original data). The histogram of linked data is colored in blue and the histogram of original data is colored in white. Therefore, the area between the white bar and blue bar represents the extent of the transactions which failed to match. After linking to the EPCs, less data was lost during the period between 2008 and 2017. Also, no particular range of house price lost significantly more as a proportion of each range (bin in the histogram) after the data linkage.


1995
1996
1997
1998
1999
2000
2001
2002
count
120000
90000
60000
30000
0
0
5
10
15
Log of house price
Data
Linked data
Original data

(A)


2003
2004
2005
2006
2007
2008
2009
2010
2011
150000
100000
50000
0
count
0
5
10
15
Log of house price
Data
Linked data
Original data

(B)



(C)

Figure 3.10 House price distribution of original data and linked-EPC Price Paid Data, 1995-2017[10]

In addition to the above visualization approach, the Kolmogorov-Smirnov test (K-S test) and the Jeffreys divergence (J-divergence) are used to further quantify the extent of house price information lost. The Kolmogorov–Smirnov (K-S) test is a nonparametric test that examines the differences in the shape of a distribution. The K-S test, statistic D, is based on the maximum absolute difference between two cumulative distribution functions. Here, the test will be used to quantify the difference

[10] Note: Original data in the graph above means georeferenced LR PPD data. Linked data means the LR PPD which can be successful link with EPC.

of two house price distributions (original data versus linked data). The Jeffreys divergences (J-divergence), derived from information theory, is a function used to establish the distance of one probability distribution to another (Jeffreys, 1946; Nielsen, 2010; Rohde, 2016). To calculate the J-divergence, the data from two different samples must first be assigned to k different categories. In the case of this research, these categories are a simple subdivision of the log house price into bins. The J-divergence is then defined as

$$J = \sum_{j=1}^{k} p^j \ln(\frac{p^j}{q^i}) + \sum_{j=1}^{k} q^j \ln(\frac{q^j}{p^i}) \quad (3.1)$$

where k is the number of categories, p^j is the proportion of data points in category j in the original house price data, and q^j is the proportion of data points in category j in the linked house price data. The final divergence measure, J, ranges from 0 to 1. If the distribution of both data samples across all the categories is the same, J will be 0. Larger values of J indicate greater differences between the two distributions.

To compute the J-divergence, the original data and linked data are divided into 150 bins. The 150 bins are created based on the 150 equal intervals of log house price in the original data in a given year. The results of J-divergence and K-S tests are shown in Figure 3.11. P-values of all the K-S tests are less than 0.05, which means there is a statistically significant difference between the original data and the linked data. The D statistic drops markedly after 2009, remaining at a low level thereafter. This demonstrates the distribution of house price before and after linkage are highly similar between 2009 and 2017. The J-divergence results also show that the final linked data exhibits relatively low information loss between 2009 and 2017. Considering the time period between 2009 to 2017, the information loss is slightly higher after 2016 than that shown by K-S. The loss of information situation after 2015 is not as bad as for the period before 2008. Both K-S test and J-divergence test shows that the newly created house price data between 2009 to 2017 is representative of the pre-linked data and can offer a more reliable dataset to represent the housing market than that for other years. As the house price data does not contain the whole of 2017, the time period 2009 to 2016 was chosen as the research period in the following analyses.

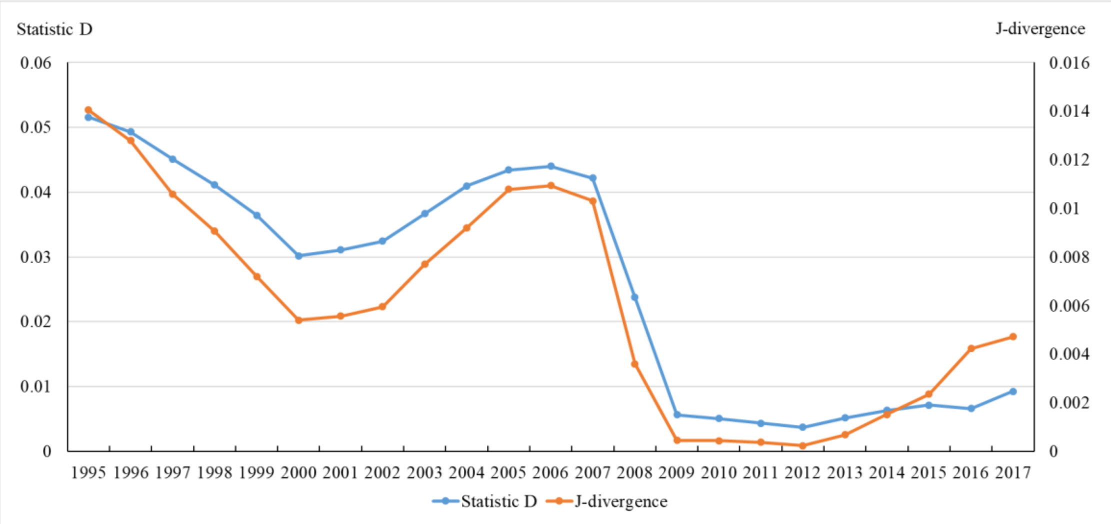


Figure 3.11 Results of K-S test and J-divergence method

3.4.2.3 Evaluation of the data linkage between 2009 and 2016

After method 1, there are 5,983,618 house price spatial data for the period between 2009 and 2016. Of these, 5,597,702 have been successfully linked to a Domestic EPC in method 2. This resulting data is named as "Linked-EPC PPD". The overall match rate for this period of method 2 is 93.55%. Table 3.6 lists the match rate by property type in this linked-EPC PPD. The match rates for detached, semi-detached or terraced houses are around 95%. However, the match rate for flats/maisonettes (89.98%) is smaller than the rates for houses (Table 3.6). This is because address elements for the flats/maisonettes are more detailed than for detached, semi-detached or terraced houses. It makes it more difficult to conduct an address-based linkage between the flats/maisonettes transactions with their domestic EPCs. The match rate for the 'Other' property type is quite small (24.34%), but this will not influence using the linked-EPC PPD to measure residential housing prices at full market value since the 'Other' category is for properties not sold at full market value[11].

Table 3.6 Summary of the matching for property type, 2009-2016

Property type	House price spatial data	Link-EPC PPD	Matching rate
Detached	1,385,966	1,309,328	94.47%
Flats/Maisonettes	1,175,397	1,057,660	89.98%
Other	47,689	11,609	24.34%
Semi-Detached	1,620,219	1,551,430	95.75%
Terraced	1,754,347	1,667,675	95.06%

The overall match rates between 2009 and 2016 by LA (Figure 3.12) are not equally distributed. The overall match rate for 95% of LAs is over 90%. The overall matching rate for the remaining 5% of LAs (17 LAs – mainly in London) is between 90% and 65%. Within these 17 LAs, the overall match rate for three LAs (City of London, Westminster and Camden) is lower than 80% (66.64%, 79.51% and 79.49% respectively). The others (Isles of Scilly, Kensington and Chelsea, Hammersmith and Fulham, Brent, Hackney, Haringey, Lambeth, Islington,

[11] Category type for 'Other' property type in LR PPD is B, which means the Other property type is not sold in the full market value. It could have transferred under a power of sale/repossessions, buy-to-lets, transfers to non-private individuals and so on (https://www.gov.uk/guidance/about-the-price-paid-data#data-excluded-from-price-paid-data). Sometimes, the Other property type is a garage rather than a real dwelling.

Brighton and Hove, Hastings, Bath and North East Somerset, City of Bristol, Wandsworth and Lewisham) have match rates between 80% and 90%.

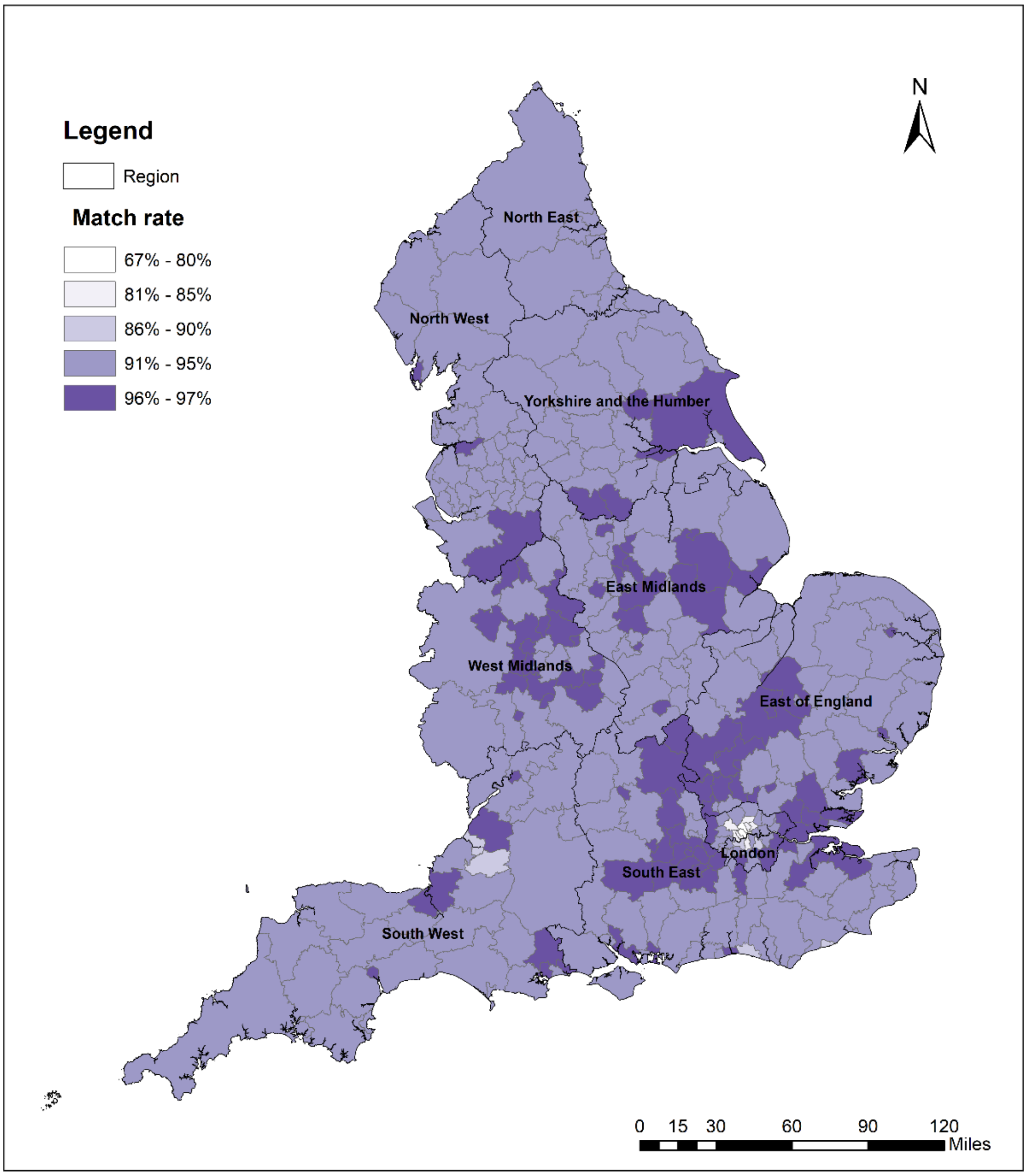


Figure 3.12 Overall match rate at LA between 2009 to 2016

Match rates in England between 2009 and 2016 are over 90% as shown in Figure 3.12. Figure 3.13 displays the annual match rate by LA. 68% of LAs have an annual match rate which is always over 90% between 2009 and 2016. The annual match rates between 2009 and 2016 are, for the majority LAs, quite stable over time with a slight fall after 2015. LAs with a high matching rate in 2009 continue with a high rate subsequently. Only two LAs (City of London and Isles of Scilly), both of which are small in terms of their numbers of transactions, show an obvious fluctuation during this eight-year period.

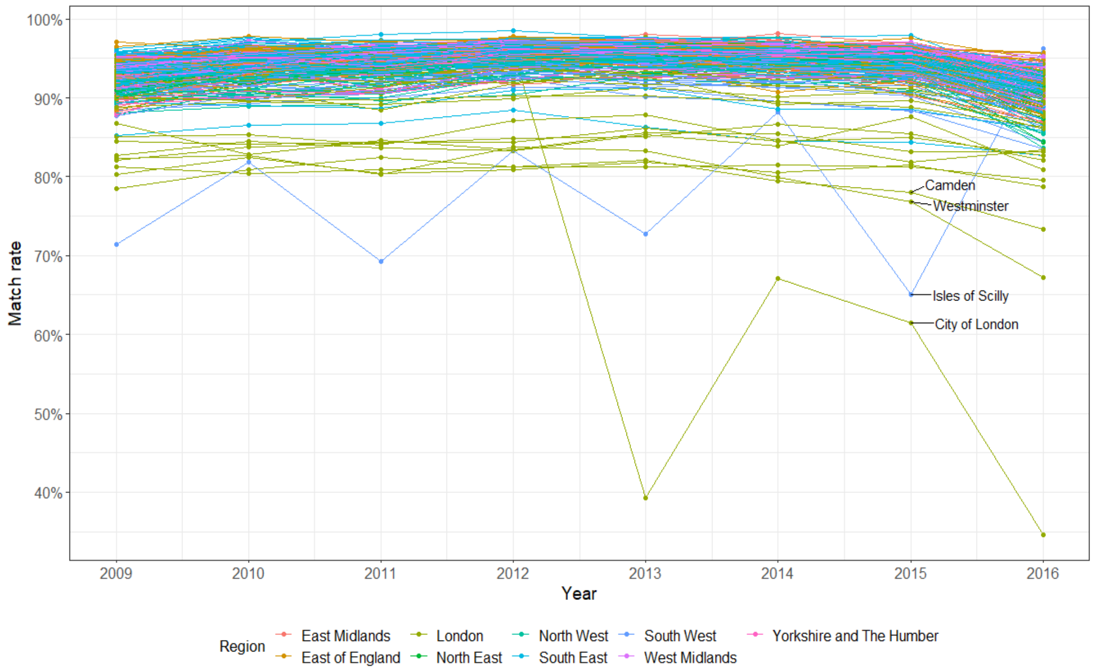


Figure 3.13 Match rate across LA in England, 2009-2016

Properties that feature in the house price spatial data (2009-2016) are not fully available in Domestic EPCs[12] (1/1/2008 -1/10/2016), which is one of the main reasons for unequal match rates across LAs. For 34,768 transactions (2009-2016) relating to 16,602 postcode units Domestic EPCs cannot be found. For example, Domestic EPCs in the City of London at postcode "EC2Y 9BA" are not available hence transactions in "EC2Y 9BA" cannot be successfully matched. In the City of London, 8.52% of

[12] Domestic EPCs are public by default, but can be withdrawn by the property owner, detail see : https://www.epcregister.com/optout. Thus properties in house price spatial data (geo-referenced LR PPD) could have an EPC but not publicly accessible.

house price transactions fail to link for this reason. Similarly, nearly 2.91% of transactions in Westminster and in Salford cannot be successfully matched. Details of the proportion of transactions at LA level with unmatched Domestic EPCs are shown for all postcodes in Appendix A4. Some transactions in house price spatial data can relate a postcode unit which is also identified in the EPC data but contain no matching property identifiers. For example, one flat sold in 2009 at Camden (Flat 65 Visage Apartments at Winchester Road) failed to match under method 2 because Domestic EPCs did not record this property.

3.4.2.4 Data cleaning

The Linked-EPC PPD comprises the transaction information in the LR PPD together with property size (total floor area and number of habitable rooms) in EPCs. Some transactions (category type B) in the LR PPD relate to property not sold at full market value. This data is excluded prior to analysis. Moreover, some properties' total floor area and number of habitable rooms are recorded in EPCs with missing or untenable values (e.g. total floor area records as 0.01). Thus, another six methods are created to clean up these outliers. All the excluded transactions along with cleaning methods are listed in Table 3.7, which in total accounts for 16.35% of the linked-EPC PPD. Missing and untenable property size values (total floor area and number of habitable rooms) in Domestic EPCs are responsible for two thirds (12.93% of all data) of the 16.35%.

Table 3.7 List of transactions exclude from the linked-EPC PPD

No.	Method	Transaction numbers	Proportion
1	Transactions where category type is B.	191,312	20.90%
2	Transactions where total floor area or number of habitable rooms are NA value or 0.	720,107	78.68%

No.	Method	Transaction numbers	Proportion
3	Transactions where total floor area is smaller than 9 m^2 or larger than 974 m^2.[13]	557	0.06%
4	Transactions where HPM is larger than 50,000 £/m^2 or HPM is smaller than 200 £/m^2.	766	0.08%
5	Transactions where floor area per habitable room is larger than 100m^2.	703	0.08%
6	Transactions where number of habitable rooms are larger than 20.	376	0.04%
7	Transactions where floor size per habitable room is smaller than 6.51m^2.[14]	1,413	0.15%
Overall		915,234	100%

After removing the transactions listed in table above, 4,682,468 transactions are ready to use for the house price analysis. This is the "enhanced house price database" shown in Figure 3.5. Method 1 (Figure 3.5) geo-references 98% of full market sales in the LR PPD in England between 2009 and 2016. Six percent of full market sales are further lost once linked with Domestic EPCs. Subsequently, 12% of full market sales are excluded owing to missing and untenable property size values in Domestic EPCs. Consequently, the Domestic EPCs' data quality in terms of property size values and data coverage are the main reason that the enhanced house price data only represents 80% of full market property sales in LR PPD in England between 2009 and 2016.

Similar to the spatial coverage of the LR PPD, the enhanced house price database fully covers all the regional areas, LAs and MSOAs in England. The LR PPD covers 99.99% of LSOAs and this is the same for the enhanced house price database. Although the

[13] According to the total floor area from the EHS (2008-2016), the range of total floor area is from 9 square metres to 974 square metres (statistics by author). All total floor area data that is not inside the range of the EHS is classified as outliers

[14] According to the min room size for one person aged over 10 years in The Licensing of Houses in Multiple Occupation (Mandatory Conditions of Licences) (England) Regulations 2018.Resources: http://www.legislation.gov.uk/ukdsi/2018/9780111167359/regulation/2

enhanced house price database is not as comprehensive as the LR PPD, it is a substantial house price dataset for England (2009-2016) containing both the TP and total floor area. It is also currently, the best accessible data for academic exploration of residential house price variation along with total floor area in England between 2009 and 2016.

3.5 Measuring housing affordability: income data review

One dimension of affordability is related to a household's housing budget. Disposable income is the most commonly used metric to determine a household's housing budget, in the UK. Table 3.8 lists a summary of open and accessible income datasets. Compared to the Living Cost and Food Survey (LCF), the Family Resource Survey (FRS) dataset has a larger sample size and longer period coverage. It is also used to calculate the number of households with below average income (HBAI) and estimate the household income at MSOA level. Furthermore, the smallest geographical level of FRS is designed at region level and is reliable when used to represent the household income distribution at regional or national level. Although FRS has quite a large sample for households, it normally uses a 3-year average method to estimate household income to avoid sampling bias. The EHS, which starts in 2008, covers home ownership, housing costs and affordability information. The EHS annual sample size is over 13,000 households and spatial coverage is across England, which is a bit smaller than FRS. However, this dataset holds information such as whether buyers are first-time buyers or not, whether they buy the property using a mortgage and when they bought the property. This information is useful to create typical households for use in the exploration of housing affordability. In addition to the above datasets, there are two non-official income statistics in Consumer Data Research Centre (CDRC). One is Individual Income Estimates (PAYE), this data is estimated at LSOA level across the England, the data divides individual gross nominal annual estimated income into 10 intervals and only offers the proportion of these 10 income categories. Currently this data is openly available but only covers 2016. It is useful to detect the household income distribution at LSOA level but not suitable to track the household income change in a given place. The other data about income from CDRC is ACXIOM - Small Area Income Data. This presents a more advanced picture of household income at each postcode in England, Wales and Scotland. This data cover only one year, 2012 and has highly restricted access. It is useful to reflect on the household income situation in 2012 at small geographical scales.

Table 3.8 A summary of data relative to household income in the England

Data / Data source	Sample size	Date Range	Smallest geographic scale	Dataset details	Data access
Family Resource Survey (FRS)	More than 20,000 households per year	1993/1994-2016/2017	Regional	Household characteristics (composition, tenure type), tenure and housing costs, income and benefit receipt	Available from UK Data Archive
Living Cost and Food Survey (LCF)	Approximately 6,000 households per year	2001/2002-2004/2005, 2006-2014, 2015/2016-2016/2017	Regional	Household expenditure, food consumption and income.	Available from UK Data Archive
Regional gross disposable household income	/	1997 to 2015	LA	/	Available from UK Data Archive
Small area model-based income estimates dataset	/	1/4/2001-31/3/2002, 1/4/2004-31/3/2005, 1/4/2007-31/3/2008, 1/4/2011-31/3/2012, 1/4/2013-31/3/2014, 1/4/2015-31/3/2016	MSOA level	Model based on a combination of data from the Family Resources Survey, the 2011 Census and a number of administrative data sources.	Open access
English Housing Survey (EHS)	Around 13,300 interview surveys;	2008/09-2016/17	Regional	Demographic, household type, duration (freehold or leasehold) housing cost (rent or mortgage payments), housing circumstances,	Available from UK Data Archive

Data / Data source	Sample size	Date Range	Smallest geographic scale	Dataset details	Data access
				, income **(**first-time buyer, owner but has owned previously, non-owner).	
CDRC Individual Income Estimates (PAYE)	100%	2016	LSOA level	10 income category proportion	Open access
ACXIOM -Small Area Income Data	100% covers for 1.2 million unit postcodes of England, Wales and Scotland	2012	Postcode level	Postcode income by age, household size	Only available on a contract basis from CDRC

Full coverage of individual household income along with house cost in England is not available. The current understanding of household income and housing costs are mainly based on survey data. This survey data is unidentified at smaller local scales (e.g. postcode) and mostly geotagged to region area information. Unlike the LR PPD, it is not possible to enhance the income dataset using the survey approach. Survey datasets have de-identified the person and address information before publication. The only choice left is based on the research purpose of choosing the most suitable survey data. All these deficiencies will hinder the understanding of housing affordability by considering the huge variety of possible household compositions. To make the problem tractable and address the research question, a suitable and achievable approach to simplify the variety of housing budgets is needed, with a consideration of current housing affordability measures. The income data, which only cover one year, are removed from the choice list as they are unable to compare housing affordability across time. The aggregate income estimate (e.g. small area model-based income estimates dataset) are also excluded as they are limit to a given geographical unit and time scale. The ratio of house price to income is the common measure in England, but this approach requires both the house price and income estimated at the same geographical level. It has high requirement for the housing budgets aspect to cover at the same geographical scales. Given this, the BBC's online housing calculator approach shows an advantage, as it only requires one household's housing budget. Furthermore, it is more achievable to create some typical households, based on the available information from surveys. The FRS data only offers the fundamentals of household income, but English Housing Survey (EHS) data offer more details of housing cost and housing affordability information underlying the different types of home ownership (e.g. weekly mortgage payment for mortgage buyer, first time buyer, non-first-time buyer). Thus, this research chooses to employ the EHS data to support the understanding of housing affordability through the ideas of the BBC calculator.

3.6 Conclusion

This chapter first outlines all the available residential house price data currently used in understanding the house price variance in England. LR PPD is identified as the most comprehensive residential house price data in terms of understanding real dynamics of housing market in England. This covers the most transaction records in England since 1995 and provides a clearer picture of house price variance compared with other house price datasets, but it limits the understanding the house price variation at small geographic scales and also by different property characteristics. Two data linkage methods are created to overcome these shortcomings to link the LR PPD, OS MMTL, OS ABP and Domestic EPCs. Although there is a certain proportion of transaction data lost during the data linkage, the new spatial attribute house price dataset shows that the date range, from 2009 to 2016, is relatively well aligned with the pre-linked the LR PPD according to the J-divergence measurement and the Chi-square test and the Kolmogorov–Smirnov test. It is relatively credible to use it in exploring the spatial and temporal house price variance. Following this, a comprehensive attribute housing price spatial database (2009-2016) is created which contains TP, property type, duration, age (old/new), dwelling total floor area, number of habitable rooms and transaction year. This enhanced dataset shows a significant value in advancing the understanding of the housing market in England at any geographical scale, plus it also enables the exploration of housing affordability issues including a consideration of property size aspects.

With a clear understanding of the house price dataset in England, this chapter moves on to exploring available income data to track housing affordability issues in England. The EHS data is useful to estimate the housing cost and household income by tenure, or household type. This will benefit the understanding of house affordability through expanding the ideas of the BBC calculator with some designated household's housing budgets estimated from the EHS.

This chapter has fully explored current available house price data and household income data in England and also builds a comprehensive database to address the thesis research questions. Understanding house price variance and identifying the most appropriate small geographic unit for the thesis research will help to advance our understanding of the spatial and temporal patterns of housing affordability. Therefore, the following chapters will be based on enhanced house price datasets to begin exploring house price variance at different geographical scales.

Chapter 4 Understanding house price variation in England: a multi-scale exploration

4.1 Introduction

Housing markets are highly geographical. However, this spatial heterogeneity in house price is often only crudely expressed at city or regional level, such as the North-South divide in England/London (Partington and Perraudin, 2018; Peachey, 2017). Understanding these variations is possible through the examination of published house price statistics. However, available house prices statistics are normally represented at a given aggregate level by the UK government and some private organisations, using various data and a range of measurement methods (Chandler and Disney, 2014; Wood, 2015). For example, the Nationwide house price index regional quarterly series is created by the Nationwide Building Society at the regional level. This is a mix-adjusted house price[15] based on Nationwide mortgage data (Nationwide Building Society, 2017). The official house price regional index, released by the ONS, also uses a mix-adjusted approach to house prices but uses LR PPD (Office for National Statistics et al., 2016). These two house price indices show the extent to which house prices differ between time periods at a regional level in England. However, they are not comparable as there are differences in data and mix-adjusted methods. House price statistics at different geographical levels that use the same dataset are still incomparable in terms of different quantity methods. For example, the ONS released two house price statistics:

---

[15] A mix-adjusted approach is an approach to statistic house price index. It is a weighted combination of house price for particular combinations of characteristics, such as location, number of bedrooms, whether or not the property has a garden or garage, and so on.

House price regional index and house price statistics for small areas (ONS, 2018; 2012). Both of these use LR PPD with two different measurements: a mix-adjusted approach and a median approach. Lack of comparability can be an issue when studying house price movements and might not be fully-understood by decision-makers; therefore, questions arise about the appropriate scale and spatial structure of aggregations, when reporting aggregated house price data.

Location greatly affects house price. From a geographical viewpoint, house price is spatially auto-correlated in small areas and also spatially heterogeneous in different geographical locations (Basu and Thibodeau, 1998; Goodman and Thibodeau, 2003; Palm, 1978). Extensive research consistently shows that the drivers behind house price variation are complex and operate at different geographical scales in the UK context (Cook, 2005; Drake, 1995; Giussani and Hadjimatheou, 1992; MacDonald and Taylor, 1993; Szumilo et al., 2017; Yao and Fotheringham, 2016). In the UK, house prices are normally analysed at regional, city, LA and sub-LA geographies such as MSOA[16]. On the broadest geographical level (e.g. regional level), house prices are influenced by macro-structural political, economic and demographic factors (Meen, 1999; Smith, 1987). On the middle geographical level (e.g. city-level), house prices are influenced by local economic conditions, local amenities, urban form and the availability of different transport modes (Downes, 2018; Smith, 2018). On the small geographical level (e.g. electoral wards), house prices are influenced by local amenities, the

[16] MSOA is a geographic area used in the Census. It is designed to support small area statistics in England and Wales, the minimum population is 5000 and the mean is 7200. There are 6791 MSOA units in England.

character of neighbouring household, local public goods (i.e. school and open space) and public transport (Orford, 2002). House prices in the same neighbourhood tend to be similar to each other, but house prices vary as a result of physical quality, such as dwelling size, age, structural design and historic value (Ahlfeldt et al., 2012; Goodman and Thibodeau, 1995; Kain and Quigley, 1970).

Meanwhile, little systematic quantitative analysis has been conducted to assess the house price variation at a range of geographical scales, particularly for the whole of England. Quantitative analysis of house price variation at multi-geographic scales have been conducted within individual cities, such as London (Feng and Jones, 2016; Law, 2018) or Cardiff (Orford, 2002; Wang et al., 2015). Except Law's study, the remaining three studies use TP to explore house price variation (Feng and Jones, 2016; Orford, 2002; Wang et al., 2015). These analyses are based on all residential transactions from LR PPD and thus they have fully explored the housing market within the individual city. Law's study started to use HPM to investigate the house price variation, at multi-geographic scales in London, but the research data (Nationwide Building Society house price) only accounted for 7% of all LR PPD (Law, 2018). It cannot fully represent the entire residential housing market and could cause problems due to the potential biases inherent in small samples (Hamnett, 1983; Jones and Bullen, 1993). Meanwhile, official house price statistics published by ONS are mainly aggregated LR PPD data, at a given administrative geography level, such as LA (ONS, 2015a) or MSOA level (ONS, 2017c).

Local variations in stock composition and other factors mean that crude aggregation to geographic units for the purposes of studying price variations is problematic. Total floor area is identified as the most important determinant of house price variation (De

Nadai and Lepri, 2018; Orford, 2010; Sirmans et al., 2006; Thwaites and Wood, 2005). To take account of the influence of total floor area on house price variation, one solution has been offered to examine house price patterns by using HPM in England, linking the LR PPD with property size (i.e. total floor area) information. Following this idea, Powell-Smith (2017) was the first to map HPM across England at postcode level. Later in the same year, ONS launched an investigation into HPM at the LA level. These two investigations advanced our understanding of house price variation at two different geographic scales, even taking into account the local variations in stock characteristics in terms of total floor area. Moreover, the enhanced house price dataset created in Chapter 3 are able to normalise TP by total floor area as HPM, to explore the house price variation.

Developing a comprehensive and systematic house price statistic for the whole of England at a variety of geographical scales will greatly aid both the government and the public in understanding housing inequality and the affordability issues that England faces (ONS, 2017f). This research aims to address this shortcoming based on the enhanced house price database described in Chapter 3 and focuses on exploring the house price variation patterns in England at multiple geographical scales. Several valid concepts of house price are currently in use: transaction house price, rent price, house price index, asking house price (Black and Diaz, 1996), on-line searching house price (Rae, 2015), HPM and house price per room (ONS, 2017b). In this research only variation of TP and HPM will be explored. The structure of this chapter is as follows. A description of study area the data is introduced in Section 2. Section 3 presents the methods used to model the house price at different geographical scales. Section 4 presents the model result and findings. Finally, this research summarizes and draws conclusions in Section 5.

4.2 Study area and data

4.2.1 Study area and geographical scales

The study area is the whole of England, one of the countries of the United Kingdom. It contains nine regions: the North East; the North West; Yorkshire and the Humber; East Midlands; West Midlands; East of England; the South East; the South Wes; and London. Administratively, England is divided into 326 LA districts, and within these there can be found 6,791 MSOA and 32,842 LSOA units frequently used for the dissemination of demographic data from the decennial Census. The extent of the study area is presented in Figure 4.1. The black lines indicate the boundaries of the regions in England. The thin grey lines indicate the boundaries of the 326 LAs of England. Regional level house prices have been well explored and shown as exhibiting a ripple effect pattern centred on London (Cook and Watson, 2016; MacDonald and Taylor, 1993). Few studies have explored house price variance at small geographical levels. Thus, this research considers three small geographical levels (from LA down to LSOA) plus the individual address level.

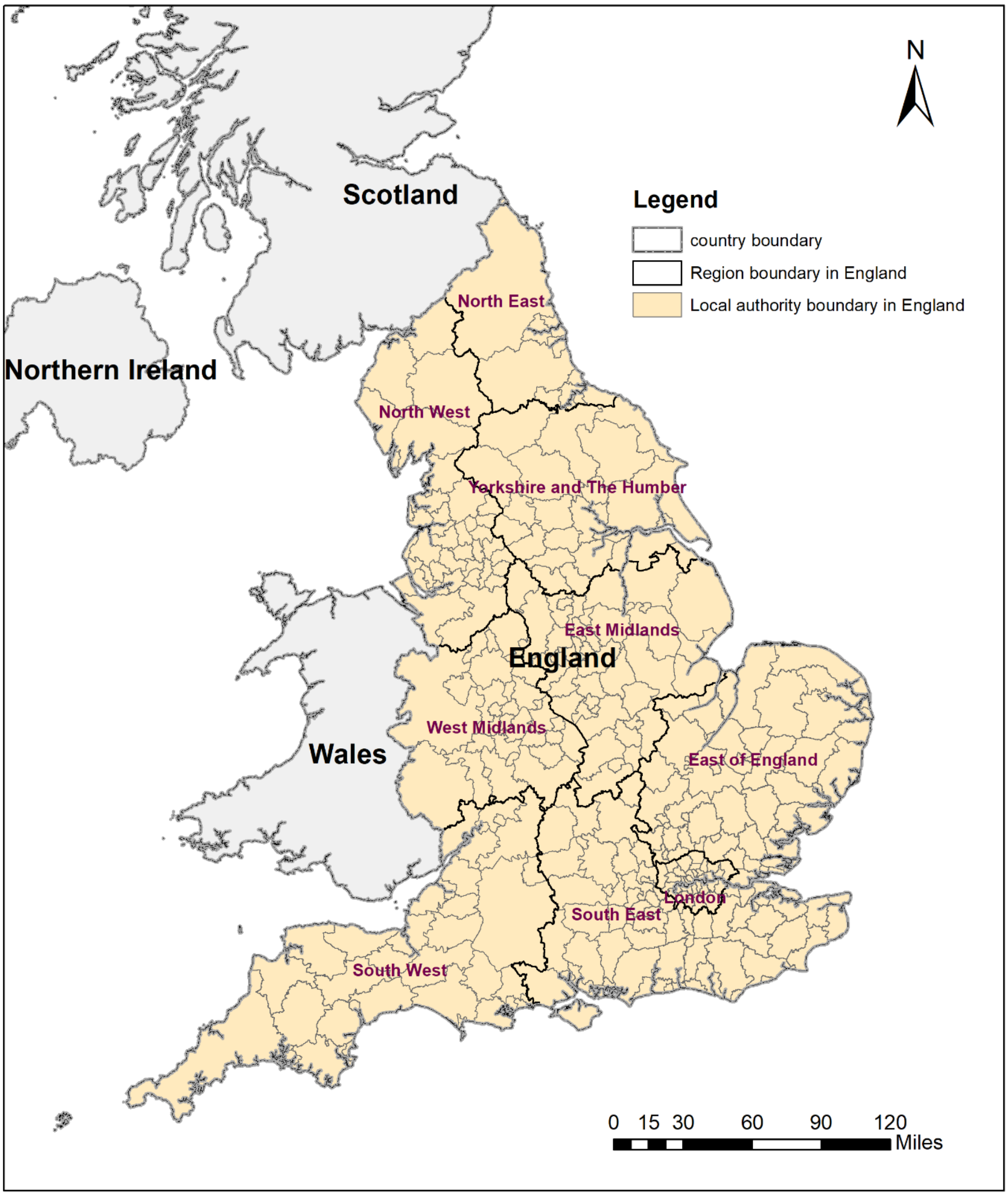


Figure 4.1 Study area

4.2.2 House price data

A newly created house price data is used to support this research, it is created to overcome the incomplete house price data issue in England and also enables the HPM calculation (Details of this creation is shown in Chapter 3). This new house price dataset contains 4,682,468 transactions across England from 2009 to 2016. Figure 4.2

shows the transaction sales for each year. Overall, there is an increasing trend in the number of transactions from 2009 to 2016. The housing market in England from 2009 to 2016 is considered an active market, since there is an increased number of transaction sales.

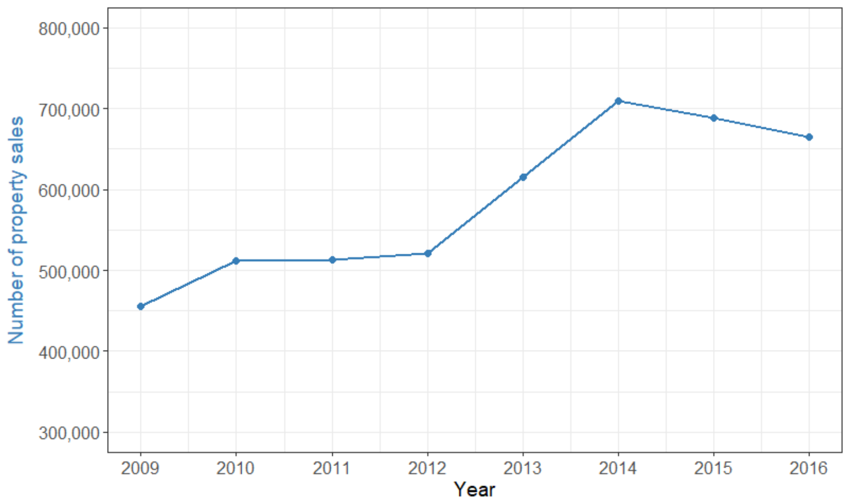


Figure 4.2 Transaction sales trend in England

Using the newly created house price database, a strong positive linear association between TP and total floor area (as measured by the Pearson correlation coefficient) can be observed within individual LAs. Figure 4.3 shows the extent of linear association between TP and total floor area, in each LA across England in 2009. For 99% of LAs, the correlation coefficient between price and total floor area (ρ) is larger than 0.5. Seventy-nine percent of LAs have ρ larger than 0.7; using the total floor area distribution in one of these LAs, 70% of the residential house price variation can be estimated. Lower correlations reveal areas where other contextual factors are having an increased influence on house prices and these can be observed in parts of London, Manchester, Liverpool and South Yorkshire.

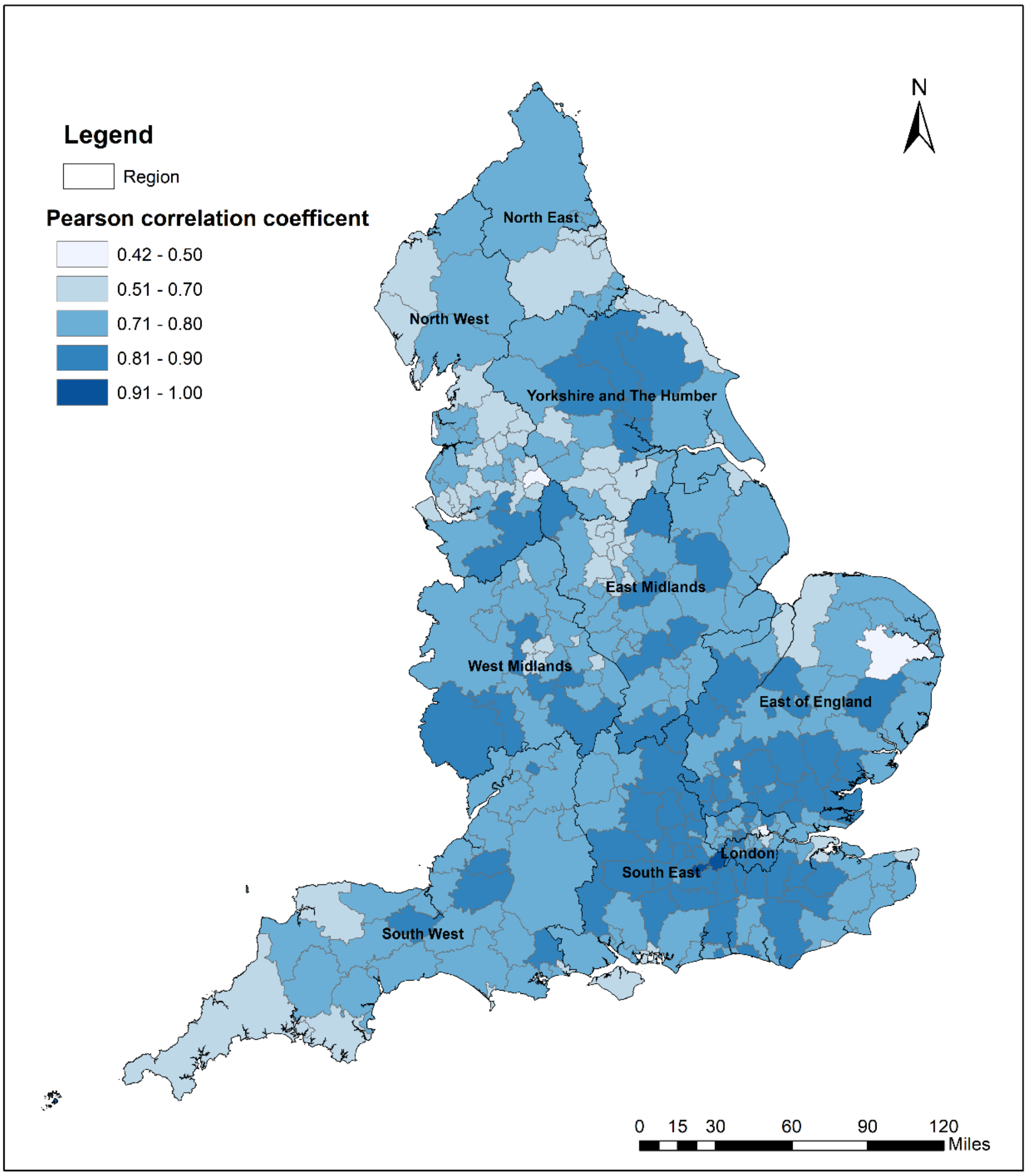


Figure 4.3 Pearson correlation coefficient at LA level in England, 2009

In some LAs, house price and total floor area show a stronger linear relationship when moved to a smaller area of analysis, such as MSOA level and property type is controlled for. One sample is shown in Figure 4.4 where, in Richmond upon Thames, local variations in floor area are particularly important for the price of semi-detached houses.

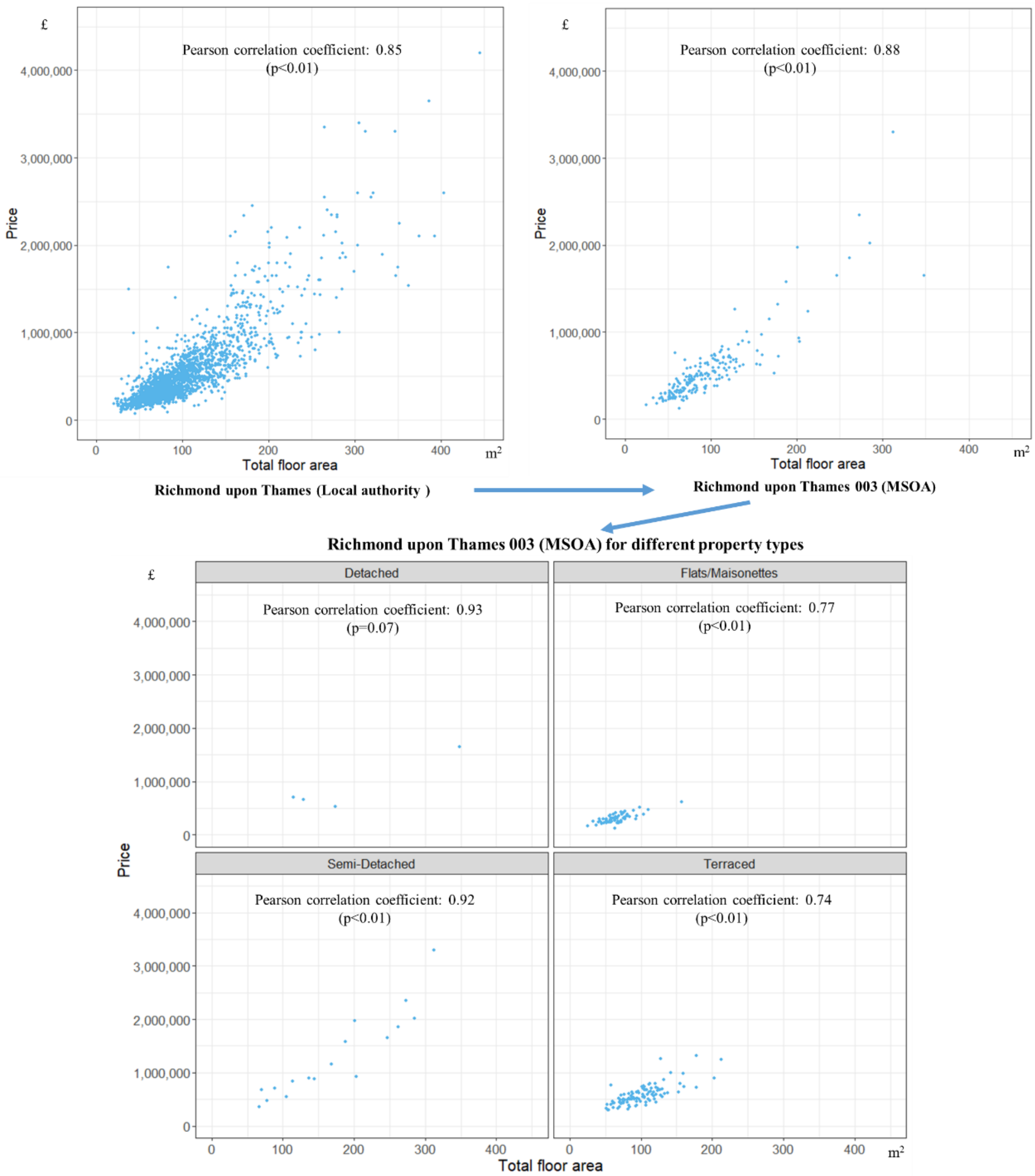


Figure 4.4 TP against total floor area in Richmond upon Thames, 2009

The high correlations between TP and total floor area are observed at LA level and MSOA level. The high correlations are also evident at some MSOAs for each property type. This high correlation observation offers evidence to consider the house price variation by accounting for the property's size effect. However, it is unclear that what is the precise difference between using TP and HPM in the house price variation exploration. Given this, the effects of using TP are compared against the more nuanced

HPM in the following analysis. Figure 4.5 plots of total TP against HPM and reveal only weak covariance. This indicates that TP and HPM do provide independent information on property prices, suggesting exploration of both is justified for this research. Seven fields in the enhanced house price database are used to support this research. They are TP, HPM, year of transaction, region codes, the 2011 Census LSOA codes, MSOA codes and LA district codes. HPM is calculated by using TP divided by the same property's total floor area.

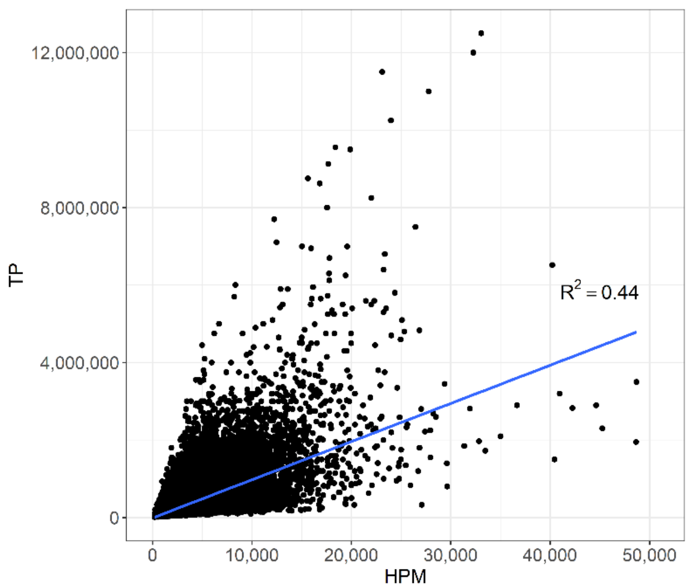


Figure 4.5 Scatter plot of TP and HPM in England, 2009

4.3 Methodology

Various modelling techniques have been proposed to capture spatial heterogeneity of prices over a large area, such as traditional hedonic price modelling (Visser et al., 2008), geographically weighted hedonic regression model (Helbich et al., 2014; Lu et al., 2014; Yu, 2007), GTWR (Fotheringham et al., 2015b; Huang et al., 2010) and MLM (Jones and Bullen, 1993, 1994; Orford, 2002). Traditional hedonic price estimation generally uses an ordinary least squares linear regression (OLS) model to identify the nature of relationships among variables. It assumes that the coefficients of the independent variables are uniform across the study area, and the error term is independently and identically distributed normally. However, house prices differ between different locations, and traditional hedonic price models ignore the variations between locations, although this problem can be ameliorated by including dummy variables for the locations. This approach has potential drawbacks when the number of locations is quite large. Geographically weighted regression (GWR) hedonic models have been developed to take into account spatial effects (Löchl and Axhausen, 2010; Lu et al., 2011) and spatial temporal hedonic models have been developed to include the additional spatio-temporal lag effects of previous sales in the vicinity of each housing sale (Fotheringham et al., 2015; Gelfand et al., 2004; Smith and Wu, 2009). However, the GWR approach has the drawback that it suffers from a heavy computational load, when dealing with large numbers of observations. Compared with the GWR hedonic models and the geographically and temporally weighted models, MLM can take any hierarchical structures present in the data into account and also have the ability to deal with more than one geographical location simultaneously (Hox, 2017; Jones, 1991a; Leyland and Groenewegen, 2003). MLM is not only able to handle big data but also results in stronger estimation through shrinkage. Therefore, MLM is the selected method in this research.

4.3.1 Multilevel variance components model

The multilevel variance components model is a multilevel model without explanatory variables. It includes only an intercept (the overall mean), the random effects at each level of hierarchy being considered, and an observation-level residual error term (Merlo et al., 2005; Ren et al., 2013). There are several synonymous names for this model, including intercept-only model, unconditional, null and empty multilevel model. In the housing context, properties can be viewed as being nested within different geographical jurisdictions. Houses are nested within neighbourhoods, neighbourhoods are nested within cities, and cities are nested within regions. Each geographical scale can be specified as a level in a multilevel model. Assuming access to a house price dataset, which records transactions (*i*) and the LAs (*j*) in which these transactions occur, the simplest multilevel variance components model, which is a two-level model could be formulated. Level 1 is the property level and level 2 is LA level. This multilevel variance components model can be written as:

$$h_{ij} = \beta_0 + l_j + e_{ij} \quad (4.1)$$

$l_j \sim N(0, \sigma_l^2)$

$e_{ij} \sim N(0, \sigma_e^2)$

Here h_{ij} is the individual house price for the *i*th transaction in geographical jurisdiction *j* (e.g. the Camden) in a given year, β_0 is the fixed effect, representing the overall mean house price, and r_j and e_{ij} are the random effects of the variance components model, representing respectively, the residuals at the LA level and the individual property level. The random effect arises from unobserved heterogeneity in characteristics that affect house prices (Feng and Jones, 2015; Snijders and Bosker, 2011). The deviation l_j measures the extent to which the mean house price in jurisdiction *j* varies from the overall mean house price (β_0), whilst e_{ij} represents the deviation of TP *i* from the mean price in its LA *j*. Residuals at the two levels (LA level and individual level) are assumed to be independent and to follow normal distributions with zero means and constant variance of σ_r^2 and σ_e^2, respectively. Moreover, residuals at the same or different levels are assumed to be uncorrelated with one another.

Figure 4.6 is a graphical illustration, which shows house price for nineteen individuals in LAs (Camden and Sheffield) in England. Individual house prices are shown as black circles, the grand mean house price (β_0) is represented by the thick black horizontal line, and the mean house prices for Camden ($\beta_0 + l_1$) and the Sheffield ($\beta_0 + l_2$) are shown as blue horizontal lines. Camden has an above average mean (positive l_j), Sheffield has a below-average mean (negative l_j). Each individual house price (i.e. h_{11}, the first transaction recorded in LA 1, Camden) is equal to the overall mean house price (β_0) plus the region-level residual for London (l_1) and its individual-level residual (e_{11}).

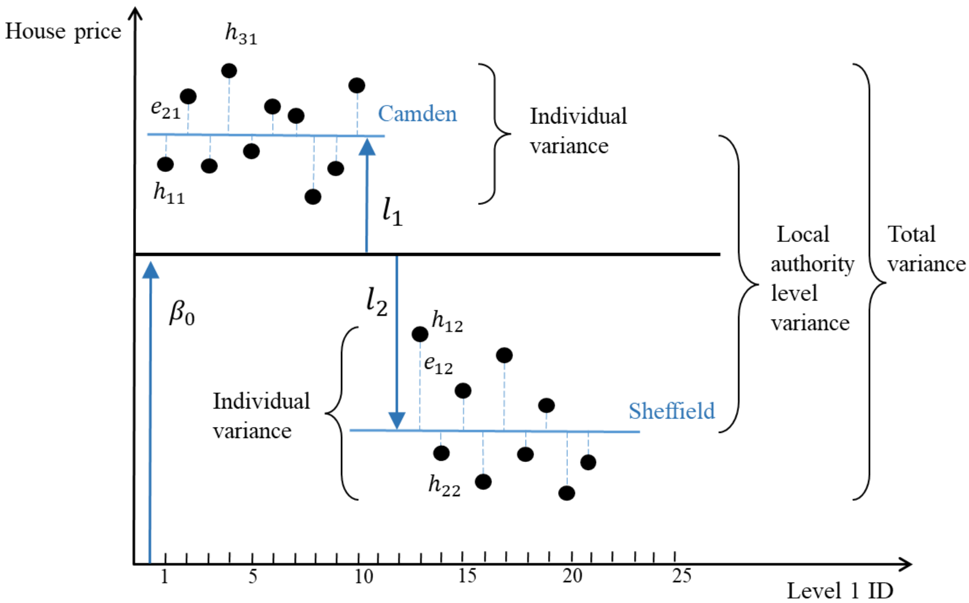


Figure 4.6 A graphical illustration of the two-level variance components model [17]

The multilevel variance components model is decomposed by the overall mean house price (fixed part) and the house price variation at each level (random part). It treats the units at each level as a random sample from a larger population with an assumed

[17] Level 1 ID in the figure stands for the identity number of each individual transaction happened in a given year.

distribution, decomposing the overall variance into the variation at two levels (σ_l^2 and σ_e^2). σ_l^2 is the variance at regional level, presenting variability among regional house prices; σ_e^2 is the residual variation at individual level. These terms are often called the "between-group variance" and the "within-group variance". Throughout this research 'group' is used as a generic term to describe clusters of individuals in terms of one specific geographical level (e.g. LA).

The two-level variance components model can be extended to three or more levels to examine the location effects at multiple scales simultaneously. Such an extension is straightforward, simply requiring the introduction of additional random effect terms. For example, a three-level model might have properties nested in MSOAs, which are nested within LAs. This is written as:

$$h_{ikj} = \beta_0 + l_j + m_{kj} + e_{ikj} \qquad (4.2)$$

$l_j \sim N(0, \sigma_l^2)$
$m_{kl} \sim N(0, \sigma_m^2)$
$e_{ikj} \sim N(0, \sigma_e^2)$

Here h_{ikj} is the individual house price for ith transaction in MSOA k of LA j, while β_0, l_j, e_{ikj}, σ_l^2 and σ_e^2 have the same meaning as before in equation 4.1. The new random term m_{kj} measures the extent to which the mean house price of MSOA k deviates from the mean house price in LA j. m_{kj} is assumed to follow a normal distribution with zero mean and constant variance σ_m^2. Residuals at the same level are uncorrelated with each other, and residuals at different levels are also uncorrelated with each other.

Variance partition coefficients (VPC) is a method for interpreting the variance components in a MLM. Taking the LA level as an example, the VPC of the LA level is calculated as the ratio of the LA level variance to the total variance. It represents the proportion of the house price variance that can be attributed to differences between regions. The VPC ranges from 0 to 1. When VPC is 0, it means there are no group differences. When VPC is 1, it means there are no within-group differences. A higher VPC at a particular level indicates that a greater proportion of total variation is due to differences between the units at that level, which indicates bigger differences between groups. In the three-level variance components model the total house price is decomposed into three variance components: individual variance (σ_e^2); MSOA level variance (σ_m^2); LA level variance (σ_l^2). The equation for VPC at individual level is presented in equation 4.3, following the equations at MSOA (equation 4.4) and LA level (equation 4.5).

Individual level $$VPC_e = \frac{\sigma_e^2}{\sigma_l^2+\sigma_m^2+\sigma_e^2} \tag{4.3}$$

MSOA level $$VPC_m = \frac{\sigma_m^2}{\sigma_l^2+\sigma_m^2+\sigma_e^2} \tag{4.4}$$

LA level $$VPC_l = \frac{\sigma_l^2}{\sigma_l^2+\sigma_m^2+\sigma_e^2} \tag{4.5}$$

The intraclass correlation coefficient (ICC) is another approach to interpreting the variance components in the multilevel model. It measures the expected correlation (similarity) of observations within groups at a particular level of the hierarchy (Bartholomew et al., 2008). This is expressed as a ratio of variances, comparing the house price variance, that occurs between groups at a particular level, to the total variation (Finch et al., 2014). This provides a measure of what is known as the cluster effect (Raudenbush and Bryk, 2002). In terms of ICC's and VPC's algebraic form, the ICC at any given level is the sum of the VPC at this level and all the higher levels. It also equates to the correlations between any two outcomes in the same level (Bartholomew et al., 2008). For example, the ICC at MSOA level is the sum of VPC at MSOA level and LA level in the three-level mode (equation 4.2). The ICC of the highest level (LA level) coincides with its VPC. The ICC ranges from 0 to 1, with a higher ICC indicating a greater degree of clustering (meaning data is more similar within groups, with larger differences between groups). Equations for the ICC from individual level to LA level are shown in equations 4.6 to 4.8:

$$ICC_e = \frac{\sigma_l^2+\sigma_m^2+\sigma_e^2}{\sigma_l^2+\sigma_m^2+\sigma_e^2} = 1 \qquad (4.6)$$

$$ICC_m = \frac{\sigma_l^2+\sigma_m^2}{\sigma_l^2+\sigma_m^2+\sigma_e^2} \qquad (4.7)$$

$$ICC_l = VPC_l = \frac{\sigma_l^2}{\sigma_r^2+\sigma_l^2+\sigma_e^2} \qquad (4.8)$$

4.3.2 Exploring spatial influences on the price variation

The multilevel variance components model was used to present an initial picture of the importance of hierarchical administrative geography units. A four-level variance components model was built to examine the clustering effect at these four different geographical scales. Level 1 is the individual residential property, level 2 is the LSOA level, level 3 is the MSOA level, level 4 is the LA level. Two sets of models, each containing eight different models, are created using this same basic structure. Set 1 is based on TPs, with separate models for each year from 2009 to 2016; Set 2 is based on HPM, again with separate models for each year from 2009 to 2016. Table 4.1 lists the details of these 16 four-level variance components models.

Table 4.1 The candidate four -level variance component models

Set	Model name	Equations
Set 1	Model TP2009	$tp2009_{ijkh} = \beta_0 + l_h + m_{kh} + o_{jkh} + e_{ijkh}$
	Model TP2010	$tp2010_{ijkh} = \beta_0 + l_h + m_{kh} + o_{jkh} + e_{ijkh}$
	Model TP2011	$tp2011_{ijkh} = \beta_0 + l_h + m_{kh} + o_{jkh} + e_{ijkh}$
	Model TP2012	$tp2012_{ijkh} = \beta_0 + l_h + m_{kh} + o_{jkh} + e_{ijkh}$
	Model TP2013	$tp2013_{ijkh} = \beta_0 + l_h + m_{kh} + o_{jkh} + e_{ijkh}$
	Model TP2014	$tp2014_{ijkh} = \beta_0 + l_h + m_{kh} + o_{jkh} + e_{ijkh}$
	Model TP2015	$tp2015_{ijkh} = \beta_0 + l_h + m_{kh} + o_{jkh} + e_{ijkh}$
	Model TP2016	$tp2016_{ijkh} = \beta_0 + l_h + m_{kh} + o_{jkh} + e_{ijkh}$
Set 2	Model HP2009	$hp2009_{ijkh} = \beta_0 + l_h + m_{kh} + o_{jkh} + e_{ijkh}$
	Model HP2010	$hp2010_{ijkh} = \beta_0 + l_h + m_{kh} + o_{jkh} + e_{ijkh}$
	Model HP2011	$hp2011_{ijkh} = \beta_0 + l_h + m_{kh} + o_{jkh} + e_{ijkh}$
	Model HP2012	$hp2012_{ijkh} = \beta_0 + l_h + m_{kh} + o_{jkh} + e_{ijkh}$
	Model HP2013	$hp2013_{ijkh} = \beta_0 + l_h + m_{kh} + o_{jkh} + e_{ijkh}$
	Model HP2014	$hp2014_{ijkh} = \beta_0 + l_h + m_{kh} + o_{jkh} + e_{ijkh}$
	Model HP2015	$hp2015_{ijkh} = \beta_0 + l_h + m_{kh} + o_{jkh} + e_{ijkh}$
	Model HP2016	$hp2016_{ijkh} = \beta_0 + l_h + m_{kh} + o_{jkh} + e_{ijkh}$
Note: tp is the TP and hp is HPM in a certain year. The parameter β_0 is the overall mean house price in England, l_h, m_{kh}, o_{jkh} and e_{ijkh} are the residuals of level 4 to 1.		

Model used in Chapter 4 is based on the TP and HPM not the log format. There are three reasons why we use price not log of price. First, VPC pattern across the four-level are the similar by using house price or log house price as outcome variable The conclusion of these research by using log model or not is the same. Second, comparing the estimated mean price and observed mean price from house price model (i.e HP2009) and its relative log house price model. The performance of these two model shows the similar result and the estimate price in house price model shows higher associate with the observed mean. Third, although the HPM is not normal distributed, Residuals at different level are symmetric and unimodal and also close to a normal distribution. This follows the basic assumption in multilevel model (variance at each level are normal distributed) Given these three reasons, this Chapter chose the most simple and easy interpretation model based on the price without log it.

4.4 Model results and discussion

Models presented in Table 4.1 were run using MLwiN 3.03 (Charlton et al., 2019). Likelihood ratio tests are used to establish whether the four-level variance components model fits the data significantly better than the null single-level model and all the related three-level models. Each four-level model in Table 4.1 is preferred to its null single-level model based on the near zero p-value of the likelihood ratio test. In addition, each four-level multilevel model was compared to a set of three-level models formed by dropping one geographic level for each comparison (e.g. dropping the LSOA level in the four-level model). All comparisons showed a significant increase in explanatory power, with increasing numbers of levels, according to the near zero p-values obtained from likelihood ratio tests. The result indicates that the test four-level models fit the data best. The following discussions are based on the estimated coefficient values for the four-level variance component models.

4.4.1 Overall house price change and house price variance

Figure 4.7.A shows the mean house price (β_0) from models in Table 4.1. The blue line represents the estimated mean TP for the eight models in Set 1, and the orange line represents the estimated mean HPM for the eight models in Set 2. Following the 2008 financial crisis, both the estimated mean TP and mean HPM show the same increasing trend from 2009 to 2016.

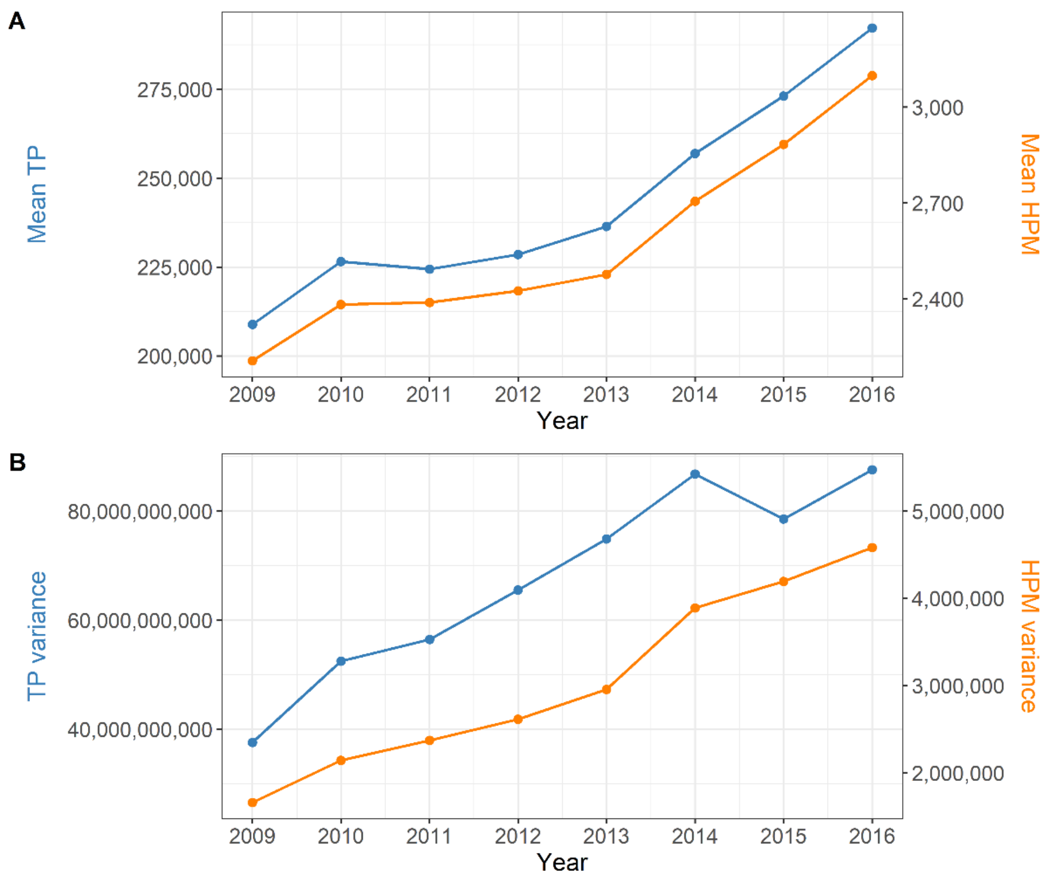


Figure 4.7 Change of overall mean house price change and house price variance between 2009 and 2016

Figure 4.7.B shows the overall house price variation trend between 2009 and 2016 based on the models in Table 4.1. Both show a trend of increase, but the trend of house price variance differs depending on whether TP is normalised by the floor area of the property – henceforth the normalised TP will be called simply the HPM, in this research. The variance of HPM increased between 2014 and 2015, while the variance of the TP decreased. Comparing the data in 2015 to 2014, a smaller number of full market value residential sales, together with fewer sales at extremely high prices, are the main reasons for the decrease in TP variance. This may be due to the increasing SDLT rates on higher bands at the end of 2014 limiting purchases of more expensive dwellings (Scanlon et al., 2017). One explanation for the trend discrepancy between

TP and HPM is the different mix of stock sold in different years. For example, a higher proportion of large dwellings (total floor area greater than 250 m^2) with high TPs (over £5 million) were sold in London's housing market in 2014, yet a lower proportion of these large, high value dwellings were sold in London in 2015 (Figure 4.8). While using the normalised price (£/m^2) approach (HPM), these large dwellings may have a low HPM; however, the small dwellings with high TPs may have a higher HPM. Therefore, the variance of HPM could increase. The overall TP variance is larger than HPM variance. This not only means that normalised TPs by the floor area are more concentrated, but also that differences in total floor area contribute greatly to TP variance. Thus, using TP, without considering the total floor area, may be misleading: the HPM is a more reliable metric for understanding house price changes.

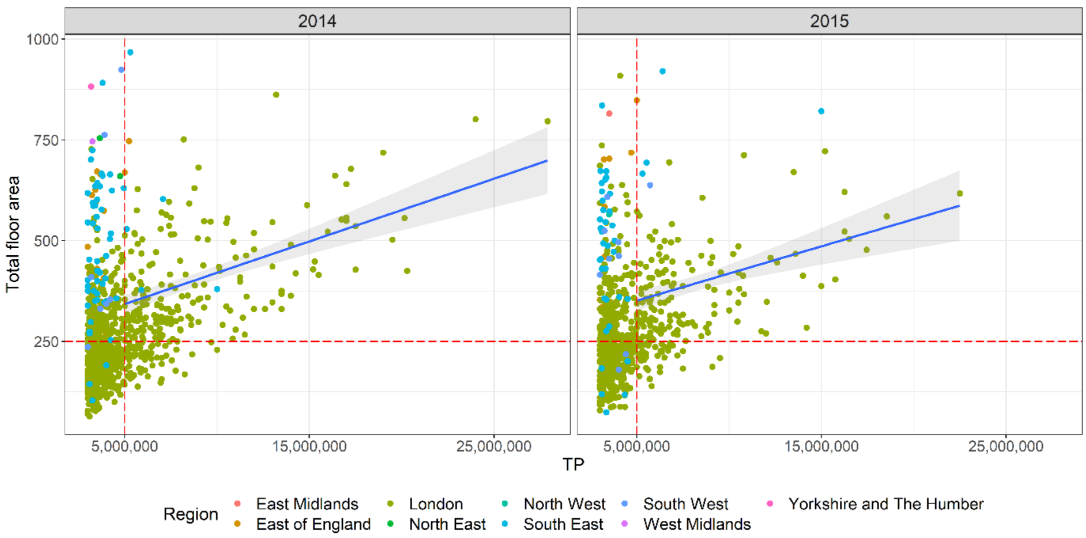


Figure 4.8 Scatter plots of TP against total floor area in 2014 and 2015

4.4.2 House price variance at four geographic scales

Table 4.2 presents the VPC and ICC results of the four-level variance components models for TP and HPM in 2009. Looking at the VPC results of Model TP2009, 44% of house price variation is caused by the three spatial effect (i.e. LA, MSOA and

LSOA). 23% of house price variation lies between LAs; 12% house price variation lies within the same LA but different MSOAs; the remaining 10% house price variation lies within-MSOAs-between-LSOAs. This means geographic scale affects house price variation: the smaller the spatial unit the greater the variability of TP.

Table 4.2 VPC and ICC statistic for Model TP2009 and HP2009

Level	Model TP2009			Model HP2009		
	Variance (million £^{2})	VPC	ICC	Variance (thousand £^{2}/m^4)	VPC	ICC
Individual level	20.95	0.56	1	551.91	0.33	1
LSOA level	3.61	0.10	0.44	65.13	0.04	0.67
MSOA level	4.41	0.12	0.35	163.13	0.10	0.63
LA level	8.62	0.23	0.23	881.18	0.53	0.53
Total	37.60	1	-	1661.35	1	-

When considering TP, the proportion of variance explained by each of the hierarchical geographic levels (ignoring the individual level) is quite small, ranging from 0.10 to 0.23. However, this pattern changes markedly when considering HPM. House price variability explained by all geographic levels (except the individual level), increases from 0.04 to 0.53. This means that controlling for floor area offsets much of the house price variation among individual properties and correspondingly the VPC changes at other geographic levels. The VPC at individual level decreases from 0.56 to 0.33. VPC at LA increases the most, from 0.23 to 0.53. This means that the most house price variation occurs at the LA level, once control is applied to the floor area effect at the individual level. The VPC at MOSA level decreases a bit, from 0.12 to 0.10. The VPC at LSOA level decreases from 0.10 to 0.04. This tiny VPC value observed at LSOA Level reveals that there is a very little house price variation occurring within-MSOA-between-LSOAs. The ICC analysis of Model HP2009 reinforces this conclusion. LA is the geographic scale that shows a big house price variation, the variation increases a bit when moving to the MSOA. However, the small variation increases when moving down to LSOA scales.

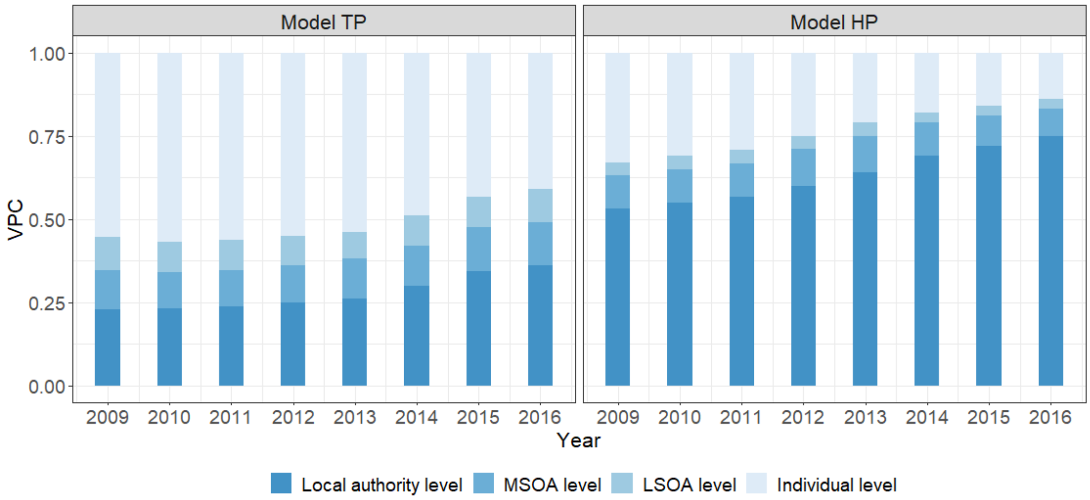


Figure 4.9 VPC results for models in Set1 and Set 2

Figure 4.9 shows the VPC results of Model TP and Model HPM for each year between 2009 and 2016. The VPC results of both set models suggest that, other factors being equal, LA effects (compared to other spatial effects, such as LSOA effects) had more of an influence on house prices in 2016 than in 2009. The proportion of TP variation (VPC), at MSOA level and LSOA level, was stable from 2009 to 2016. In contrast, the VPC at LA level (represented by the dark blue bar) doubled from 2009 to 2016. This means the LA house price differences became greater over this period. HPM variance across different levels follows a similar pattern to TP variance, but with a higher variability at the same level. Comparing VPC at the same level and same year for TP against HPM, reveals that once variations in property size are controlled for, spatial effects become much stronger. This reveals that controlling for floor area offsets much of the house price variation among individual properties and correspondingly increases VPC at higher geographic levels (i.e. level 4 to level 2). LA effects (compared to other two spatial effects) had more of an influence on TP in 2016 than in 2009, but when floor area is accounted for, this change in the LA's influence is even more noticeable. This suggests that analyzing house price difference, without considering the property size (i.e. total floor area) difference, at these three levels will hold back the understanding of house price variation. It is clear that the HPM offers a better insight into house price differences. Thus, the remainder of this research is based on the HPM model in its quest to understand house price variation in England between 2009 and 2016.

4.4.3 HPM clustering at four geographic levels between 2009 to 2016

The ICC, at a given hierarchical level, shows similarities with the HPM between administrative units at that level. Thus, ICC offers the degree of house price clustering at the given level. The ICC results of set 2 models are presented in Table 4.3. ICC at LA level is 0.53 in 2009 and continues increasing to 0.75 by 2016, illustrating that HPM are clustered at LA level. ICC at MSOA level is 0.63 in 2009 and continues increasing to 0.83 by 2016. Meanwhile, the ICC at MSOA level shows negligible improvement at LSOA level. This suggests that HPM at MSOA level are highly clustered and variations within the same MSOA unit are quite small between 2009 and 2016. This also suggests that using the mean HPM at MSOA level gives a relatively clear house price picture (2009-2016) and very little additional explanatory power is gained from observing house price variations at a more granular geographical scale. This spatial association is helpful, as this highly auto-correlated relationship between the HPM, at the MSOA level, makes predicting house prices at this level more reliable.

Table 4.3 ICC results for multilevel models in Set 2

Level	Model HP							
	2009	2010	2011	2012	2013	2014	2015	2016
LA level	0.53	0.55	0.56	0.60	0.64	0.69	0.72	0.75
MSOA level	0.63	0.65	0.67	0.72	0.75	0.79	0.81	0.83
LSOA level	0.67	0.69	0.71	0.75	0.79	0.82	0.84	0.86
Individual level	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

The ICC at MSOA level is equal to the VPC at MSOA level plus LA level. Owing to a noticeable VPC increase at LA level between 2009 and 2016, ICC at MSOA level shows a strong increase. HPM became more highly clustered at MSOA level between 2009 and 2016 as shown by the increase in ICC, which is largely due to the noticeable VPC increase at LA level.

4.4.4 Exploring house price variation at LA level

Owing to the total HPM variance increases between 2009 and 2016 (Figure 4.7), increasing house price variance at LA level is the main reason behind the VPC increase between 2009 and 2016 (Figure 4.9). Figure 4.10 shows estimated residuals at LA level (l_j) from set 2 models as scatter plots. Each point represents the residual of one LA and the same-coloured points belong to the same region. Residuals are ranked across England. The red horizontal line is the zero residual line, which presents the overall mean house price in England (β_0). Points above the line represent LAs with a mean HPM greater than the overall mean, and those below the line represent LAs with a mean less than the overall mean. It is obvious that the house price variance at LA level is largely due to some LAs in London with extremely high house prices.

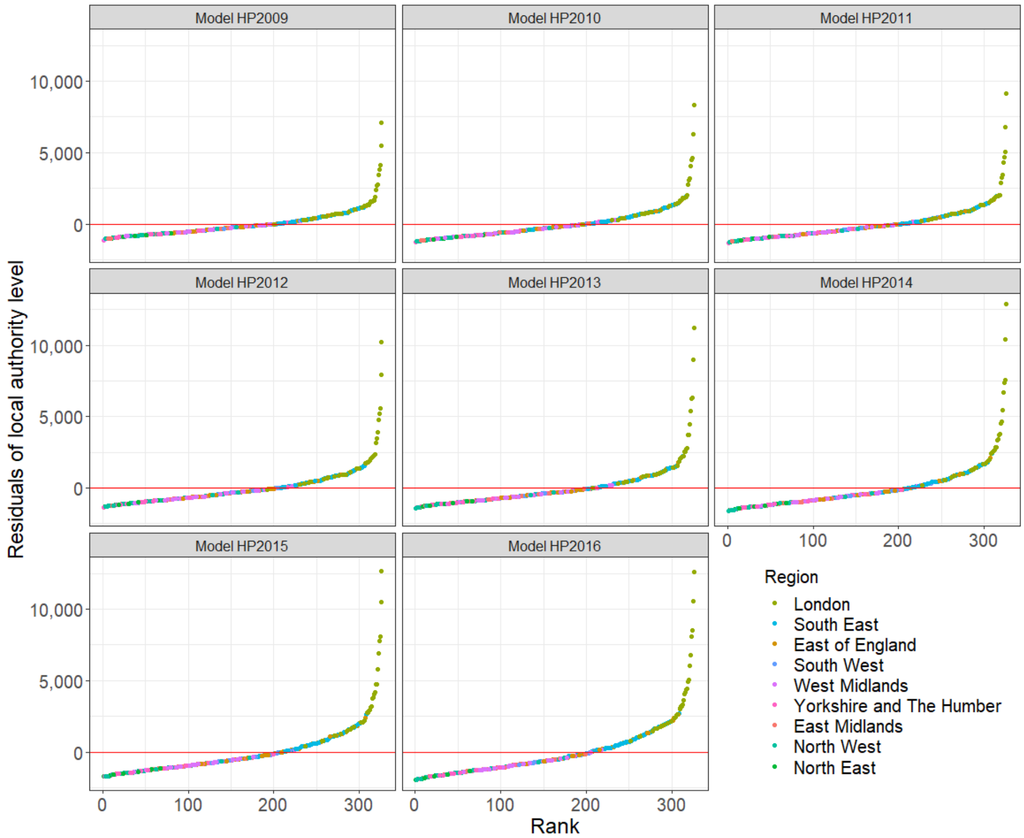


Figure 4.10 Residuals at LA level in England for models Set 2

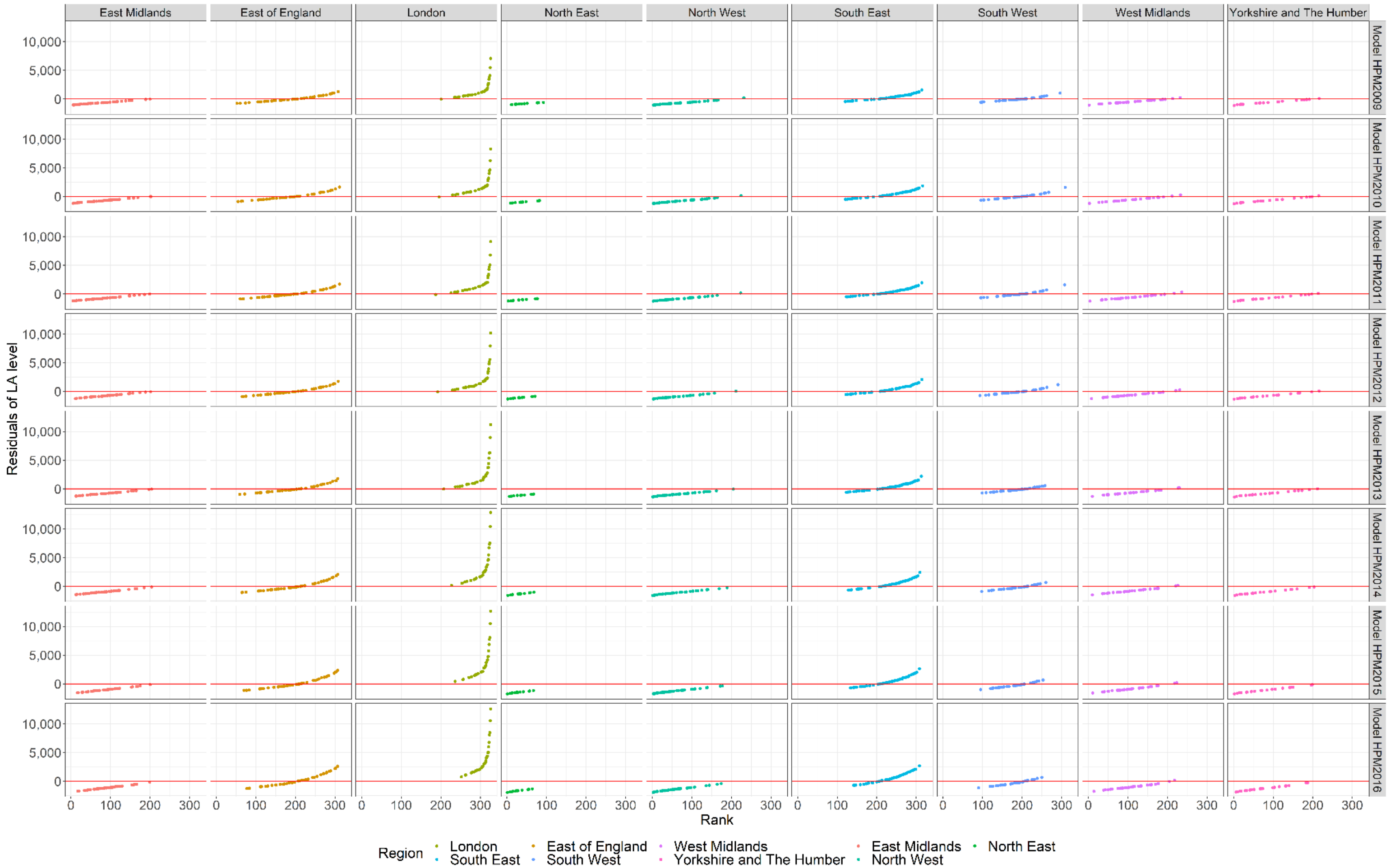


Figure 4.11 Residuals at LA level for the models in Set 2

Examining LA residuals further by plotting them for different regions (Figure 4.11), shows that, with the exception of Barking and Dagenham, house prices in each of London's LAs are consistently above the overall mean house price in England, with a continuously widening house price difference. London can be classed as an 'outlier' region in England and maintains its position as the most expensive region. London's LAs display a more rapidly increasing house price than the LAs in other regions. This London effect dominates the increasing house price variation at LA level from 2009 to 2016. Meanwhile, relatively small house price increases in the North East and the East of England also make a small contribution to the widening differential in LA's house prices.

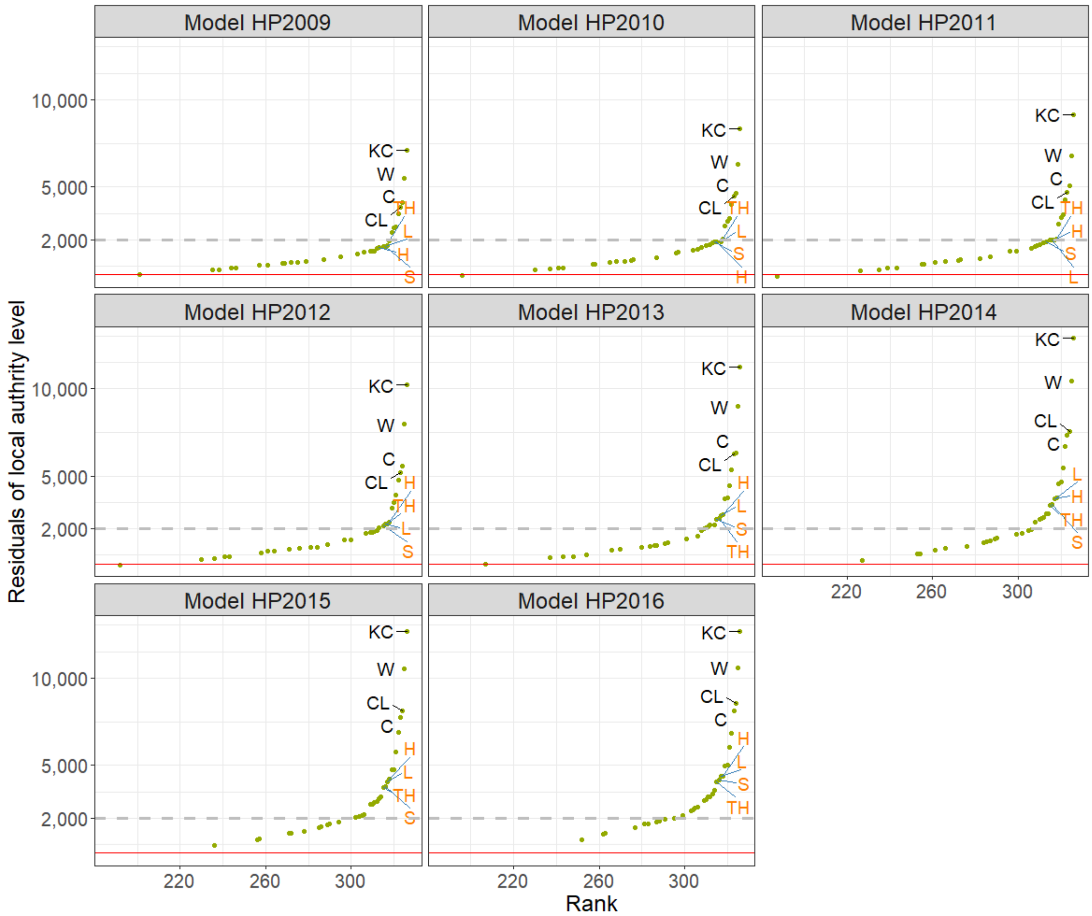


Figure 4.12 Residuals at LA level in London for the models in Set 2

Figure 4.12 displays LA residuals in London (third column plot in Figure 4.11) and offers a graphical illustration of London's increasing house price variation from 2009 to 2016. Underlying the context of rapid increases in London's house prices, some 'outliers' LAs with extremely high house prices dominate and provide the main contribution to increasing mean house price in London. The borough of Kensington and Chelsea (KC) consistently tops in this contribution, followed by Westminster (W) and then Camden (C) and City of London (CL). Some other LAs also show a substantial increase in house prices between 2009 and 2016, contributing to the increasing LA house price difference. For example, Southwark (S), Lambeth (L), Hackney (H) and Tower Hamlets (HM) were lower than the 2,000 level (grey dashed line) from 2009 to 2011, but after that, their increasing prices started to exceed the 2,000 level. As a result of these change, the house price pattern at LA level in London formed a clear rich-poor divide pattern. Eleven LAs belonging to this high house price part are Kensington and Chelsea, Westminster, Camden, City of London, Hammersmith and Fulham, Islington, Richmond upon Thames, Wandsworth, Tower Hamlets, Lambeth and Southwark.

To further explore Figure 4.12, Figure 4.13 displays the LAs' residual rank change in London from 2009 to 2016. The number inside each solid circle gives the residuals rank number within London. Smaller numbers indicate higher house prices. Looking at the rank pattern of residuals within London, Kensington and Chelsea consistently tops the price league. Moreover, six LAs continue to have the relatively highest house prices, they are Kensington and Chelsea, Westminster, Camden, City of London, Hammersmith and Fulham, and Islington. Meanwhile, Barking and Dagenham consistently exhibit lowest prices. Ten LAs (with green colour) display an increased rank order from 2009 to 2016. The remaining seventeen LAs generally display a decreased rank order.

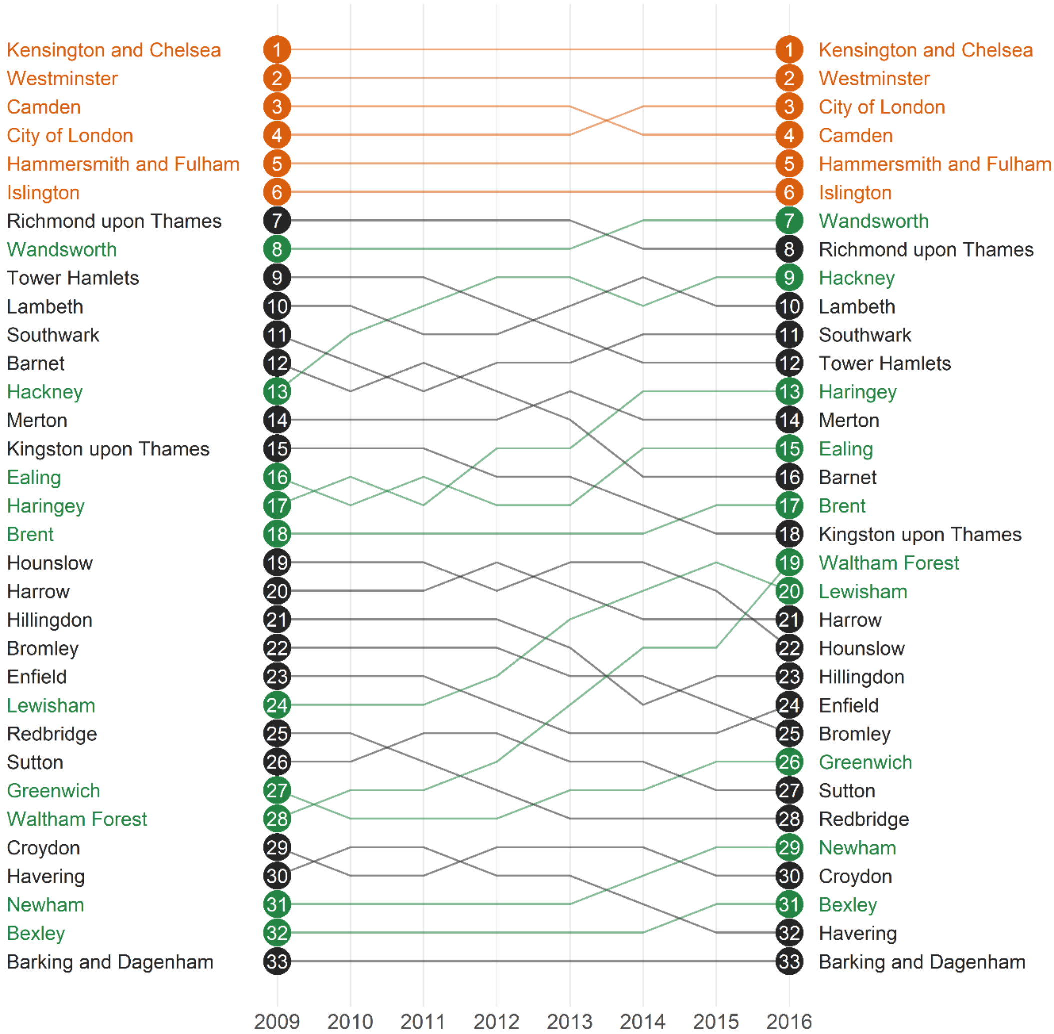


Figure 4.13 Ranks of LA's residual from 2009 to 2016

Figure 4.14 is a geographical version of LA's residual in 2009. The residuals are classified into six groups based on the natural breaks method. The darker shading indicates higher HPM residuals and each LA's name is coloured the same as in Figure 4.13 to represent the LA's rank order change. The London housing market follows a gradient pattern at LA level, with Kensington and Chelsea and Westminster at the centre, decreasing in price away from the centre. In addition, examining both Figures 4.12 and 4.14, four LAs (Southwark, Lambeth, Hackney and Tower Hamlets) show a rapid rise after 2013. These are all located in the east of London and are bordered by most expensive LAs (such as Westminster, Camden, Kensington and Chelsea and Westminster). This reveals that, after 2013, the house prices in the eastern central group of LAs (Southwark, Lambeth, Hackney and Tower Hamlets) draw away from outer London prices and closer to those of the central and South Western LAs. The LAs labelled with green text in Figure 4.14 are the LAs showing increases in their rank order from 2009 to 2016 in Figure 4.13. Locating their position in Figure 4.14, it is clear that these LAs almost form a ring around Kensington and Chelsea, Westminster, Camden, Islington, Hackney, Tower Hamlets, Suothwark and City of London). They mainly located in the east and southwest of centre London. This again reveals in London, the trend of highest house prices being in Kensington and Chelsea and Westminster has spread out into the southwest and east part of London.

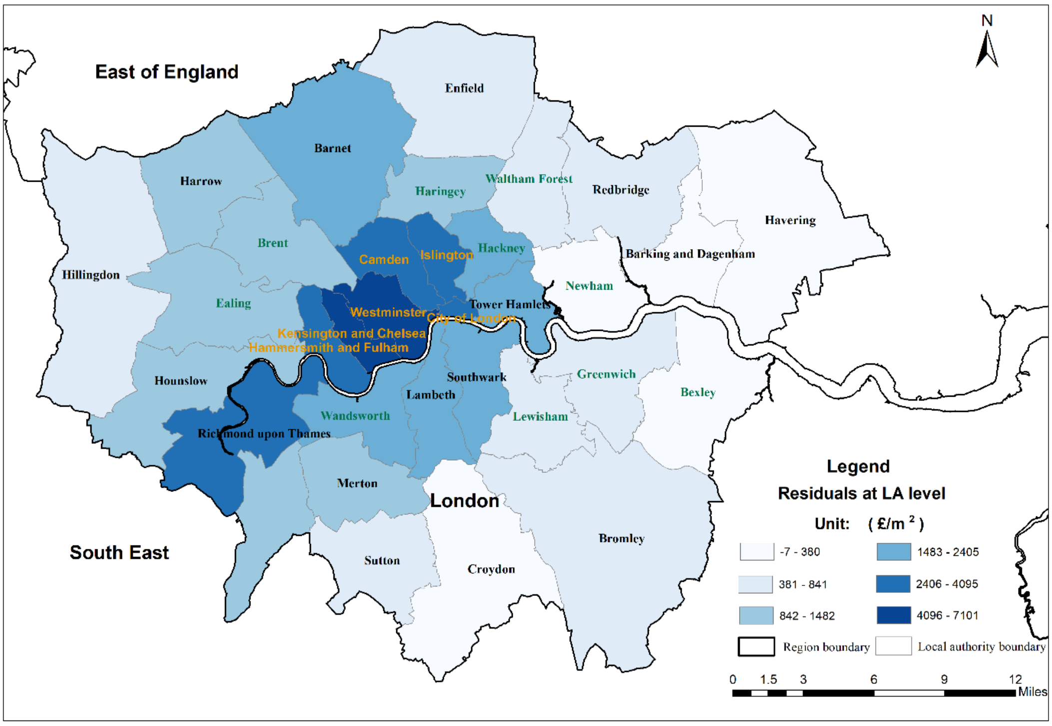


Figure 4.14 Residuals at LA level in London, 2009

London's high house price not only influences its internal housing market but also affects neighbouring housing markets. Figure 4.15 displays the residuals at LA level in England at 2009 with a separate map of housing markets surrounding London. Residuals at LA level in England are grouped into eight sub-groups using the natural breaks method. The LAs with darker shades show relatively high house prices when compared to the price in England. The LAs with the highest relative house price in the South East and the East of England are all located near to London. The eleven LAs labelled in the zoomed map of housing markets near to London are all in the top 30 price league in England in 2009. Furthermore, London's effect on the South East is stronger than on the East of England, as the proportion of LAs with high house prices is greater in the South East than in the East of England. This could be due to the higher density of commuter rail routes to the south west of the London when compared to the north east of the capital.

Setting aside London's effect within its housing market and on the housing markets in nearby LAs, the remaining eight regions still show variability at LA level. The LAs with a darker shade indicate those with a high house price relative to the mean house price in England (β_0). Excepting LAs with the highest HPM that are located near London in the South East and East of England, Cambridge exhibits a high house price in the East of England. In the South West, LAs near the southern coast show the highest house price. These are South Hams, Purbeck and East Dorset. Meanwhile, Cotswold in the north east part of the South West region also shows a high HPM. Near to the Cotswolds, a couple of LAs in the West Midlands, the South East and the East Midlands show a higher house price compared to their regional house price; these are Warwick and Stratford-upon-Avon in the West Midlands, and Daventry and South Northamptonshire in the East Midlands. Moving to the North of England, a group of contiguous LAs near to large national parks exhibit a relatively high house price. These are Eden, South Lakeland, Richmondshire, Hambleton, Ryedale, Craven, Harrogate, York and Ribble Valley. These LAs with their high house prices are the reason for the big difference in house price differentials at LA level.

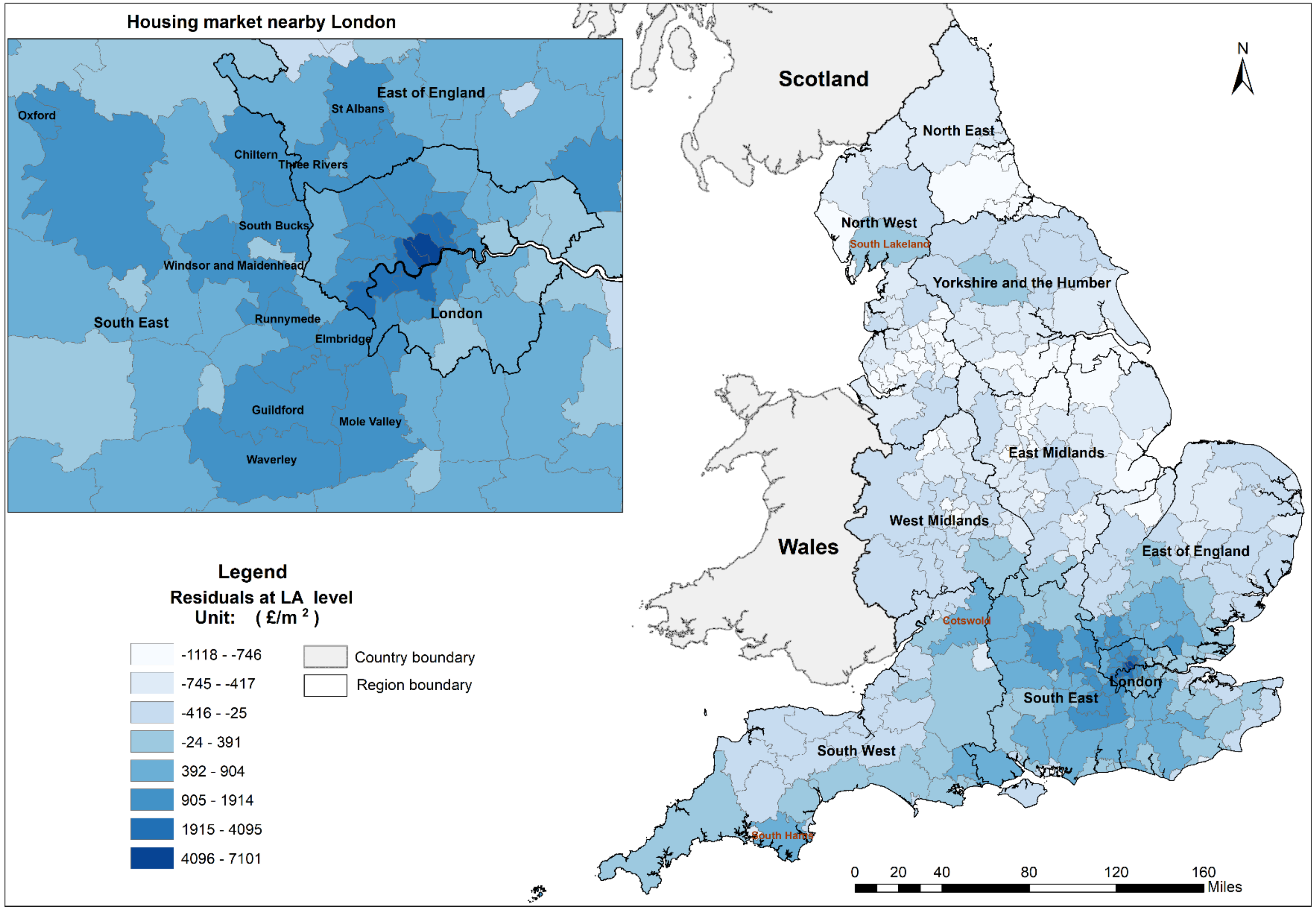


Figure 4.15 HPM residuals at LA level in England, 2009

4.5 Conclusion

Understanding the nature of differentials in house prices at different geographical levels leads to a better understanding of the housing markets. This study contributes to house price research in three main ways: Firstly, it compares house price variation in TP and HPM across England through the use of newly created house price data. This new house price dataset is based on LR PPD, which contain useful attribute information such as total floor area, and the geospatial location (geocode) of each property. Secondly, this work confirms that HPM offers a more detailed picture of house price variation than TP alone. This new approach changes the house price variability picture considerably, demonstrating floor area to be an important factor, which should not be disregarded. To better reflect house price variation, the HPM metric should be used. The third and most central contribution of this work is its examination of house price variation at four geographic scales (LA, MSOA, LOSA and individual) across England using a four-level variance components model. This has shown that different geographical scales exhibit differential spatial impacts upon house prices and that these impacts changed between 2009 and 2016. The LA effect on house price is quite evident, as is the spatial effect increase at MSOA level. The spatial effect on house price variation change is very small when examining the data at LSOA level. HPM are shown to cluster at the LA level and to be highly clustered at MSOA level. HPM differentials between LAs are quite large and this differential continues to increase a little at MSOA level. Within the same MSOA unit in England, HPMs are very similar. The LA effect on house price is important and the spatial effect further increases – again modestly – at MSOA level.

Overall house price variability in England shows an increase from 2009 to 2016. In

2009, 53% of house price variation existed between LAs. The magnitude of disparities increased 1.42 times in the following eight years. In 2016, 75% of house price variation existed between LAs. Within LA between MSOAs, there was a further 10% variation in HPM. It is clear that there is a major variation in house price occurring between LAs, rather than between MSOAs within LA. As housing regulation and delivery in England has been carried out by LAs (LAs), the conclusion is drawn that LA level is the most appropriate geographical scale at which to start to understand house price variation, with the MSOA level being the second-most appropriate scale.

Looking at HPM variation between LAs between 2009 and 2016, through plotting the residuals of LA level, we found that some LAs in the central part of London are the main reasons for this increasing LA effect. Moreover, London affects not only house price differences between regions but also its nearby LAs. LAs in the South East and East of England, which are near to London, show the highest house price within their regions. Of the top 30 LAs with the most expensive house prices, ten LAs are located outside London. The current housing policy (e.g. Right to Buy, Help to Buy) differentiates between London and outside, it would be more consistent to base the policy on house prices, that include some of the expensive areas bordering London. This would improve the consistency of the policy, although not necessary its suitability.

This research has also demonstrated that multilevel variance components modelling can offer an efficient and systematic measure for the exploration of house price variation at multi-geographic scales and provides a new insight into spatial disparities in house prices across England. This methodology can be adopted in other countries across the world to help the policy maker or public have a better understanding of house price differences and to identify the housing markets with extremely high prices.

Going beyond spatial effects, time is another determinant of house price variation that needs to be explored more deeply. Therefore, the next chapter extends this work to have a clear understanding of time effects on the house price variation, with control of the LA and MSOA spatial effect. Understanding the mechanism of house price variation across space and time will not only offer deeper insights into pressing housing inequality issues in England, but also better assist the understanding of the spatial and temporal patterns of housing affordability.

Chapter 5 Delineating the spatio temporal pattern of house price variation by local authority in England: 2009 to 2016

5.1 Introduction

A house is an immovable asset and its location is regarded as the most important determinant of its value (Downes, 2018; Kiel and Zabel, 2008). However, this house price heterogeneity in the UK is often crudely expressed as the "North-South divide". A more finer house price heterogeneity has been well-explored at regional house price level and conceptualized as a ripple effect, with London and the South East playing a leading role in terms of spill overs to other regions since the 1980s (Alexander and Barrow, 1994; Cook and Watson, 2016; Giussani and Hadjimatheou, 1991; MacDonald and Taylor, 1993). England's house prices exhibit large spatial disparities at region level since 1969 with an increasing widening of regional disparities after 2009 (Hamnett and Reades, 2019). The regional house price spatio-temporal patterns and fluctuations have been explored in some detail (Cook, 2003; Cook and Watson, 2016; Hamnett and Reades, 2019; Meen, 1999; Stevenson, 2004), but little sub-regional analysis has appeared in the literature (Cooper et al., 2013; Gray, 2012). Up to now, no research has focused on house price spatial and temporal patterns at, or lower, than LA level across England after 2008, which was the time of the GFC – a time of great shock in the English housing system. This research aims to fill this knowledge gap by exploring the housing price in England at and below LA scale, with a focus on the period after the GFC.

England's HPM are found to be clustered at LA level and highly clustered at MSOA level in Chapter 4, but gaps in our understanding still exist where the recent interacting influences of space and time on HPM are not fully understood. Therefore, this chapter set out to overcome these shortcomings and further explore HPM variation in England at and below LA level, across different temporal scales, offering new observations on price variations across space and time. The particular time scales are chosen because they are commonly used time slices in analysis of house price trend in England. Previous research on regional house price trends in the England has used data aggregated by quarter or by year (Ashworth and Parker, 1997; Gray, 2012; Hamnett and Reades, 2019). Meanwhile, a few house price analyses in England chose a half year time scale (Osborne and Neate, 2021; Vincent, 2020).Thus, these three different time scales are chosen to be explored in this research. The aims are twofold – firstly, to understand the extent to which space (LAs and MSOAs) and time (year, half-year and quarter) influence house price variation in England. Secondly, to facilitate a deeper understanding of spatio-temporal patterns of house price change at LA level in England using GCM and hybrid hierarchical k-means clustering. The structure of this chapter is as follows. Section 2 provides a description of national and regional house price change. Section 3 presents the variance components model, the GCM and hybrid hierarchical k-means clustering method used to understand the spatial and temporal pattern of house price variation. Section 4 presents the research results. Section 5, summarises the chapter and draws initial conclusions, together with recommendations for future research.

5.2 Research Data

5.2.1 House price data

Using the newly created house price data from Chapter 3 the incomplete house price data issue in England is now to be addressed. The new dataset was created by linking the LR PPD, OS MMTL, OS ABP and Domestic EPCs through two complex address-based data linkage methods. This new house price dataset records 4,682,468 transactions in England from 2009 to 2016. It records HPM by different time and location in England between 2009 and 2016. Figure 5.1 displays the HPM density plots below 15,000 £/m^2 in England between 2009 and 2016. HPM distributions in each year are seen to have a positively skewed house price but close to a normal distribution over this period.

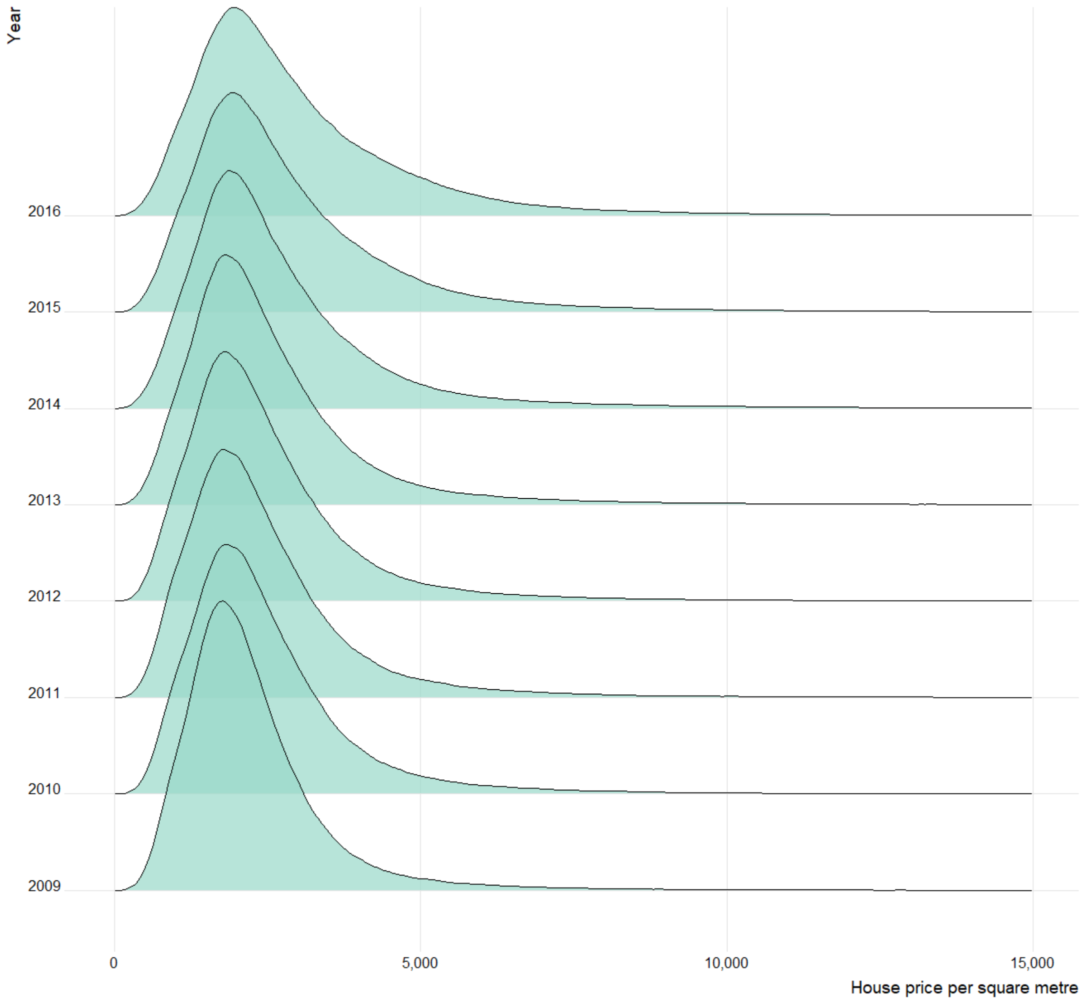


Figure 5.1 HPM density plots in England

5.2.2 National and regional HPM trend in England since 2009

The annual average HPM at national and regional levels in England between 2009 and 2016 are shown in Figure 5.2. The house price in England (the black line) shows a continuous upwards trend. The average house price in England is 2,270 £/m^2 in 2009 and it increased by 32% between 2009 and 2016. HPM trajectories at regional level also show an increasing trend but differ in terms of overall growth rate. London shows the greatest overall growth (62%) between 2009 and 2016, followed by the East of England (44 %) and the South East (42%). The remaining six regions in England have an overall increase rate below the average rate for England. The South West, East Midlands and West Midlands have increase rates of 27%, 25% and 21% respectively. Yorkshire and The Humber and the North West have an increase rate near 13%. The overall increase rate of the North East is quite small at only 3%.

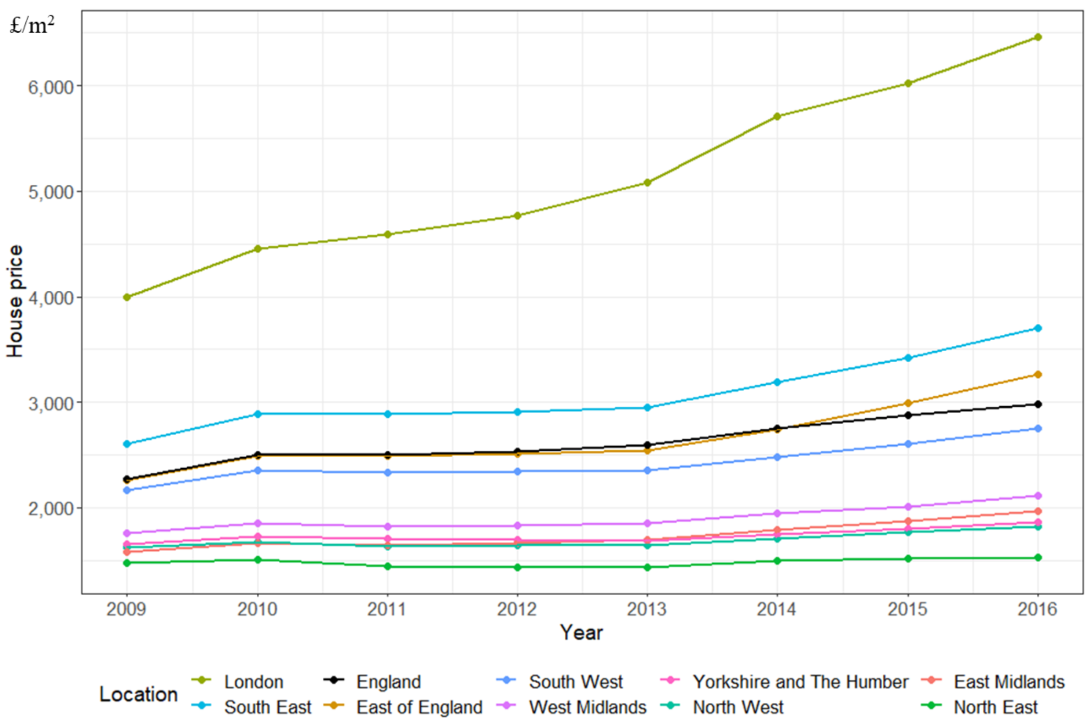


Figure 5.2 HPM trends at national and regional levels in England

Inter-regional average HPM differentials are substantial and this inequality intensifies between 2009 and 2016. London is a stand-alone housing market whose changing patterns are distinctly different from the other English regions. Mean house price for the most expensive region (London) is 2.71 times that of the mean house price for the cheapest region (North East) in 2009. This ratio increases to 4.23 in 2016. Figure 5.2 also displays three clear groups. London is in group 1, with its exceptionally high prices (3,998 £/m^2) and the fastest growth rate (62%) over the eight-year period. The East of England, the South West and the South East are in group 2, with high house prices (around 2,300 £/m^2) and a rate of increase around 37%. The remaining five regions comprise the third group, which has low house prices (around 1,600 £/m^2) and a generally slow increase rate of around 15%. Overall, the regional house price patterns in England show the North-South Divide pattern with growing differentials.

5.3.3 HPM trends at LAs in England

The annual average HPM at LA level in England between 2009 and 2016 are shown in Figure 5.3. The LA house prices display a clear non-linear growth, as do the national and regional house prices. The trends vary between different LAs. The LAs (e.g. Kensington and Chelsea) with high house prices in 2009 tend to have high rates of increase. Owing to increasing SDLT rates on higher bands at the end of 2014, which limited purchases of more expensive dwellings (Scanlon et al., 2017), the increase trend for the most expensive LAs (labelled in Figure 5.3) slowed after 2014. House price differentials across LAs increased between 2009 and 2016, and this difference is greater than the inter-regional house price differential.

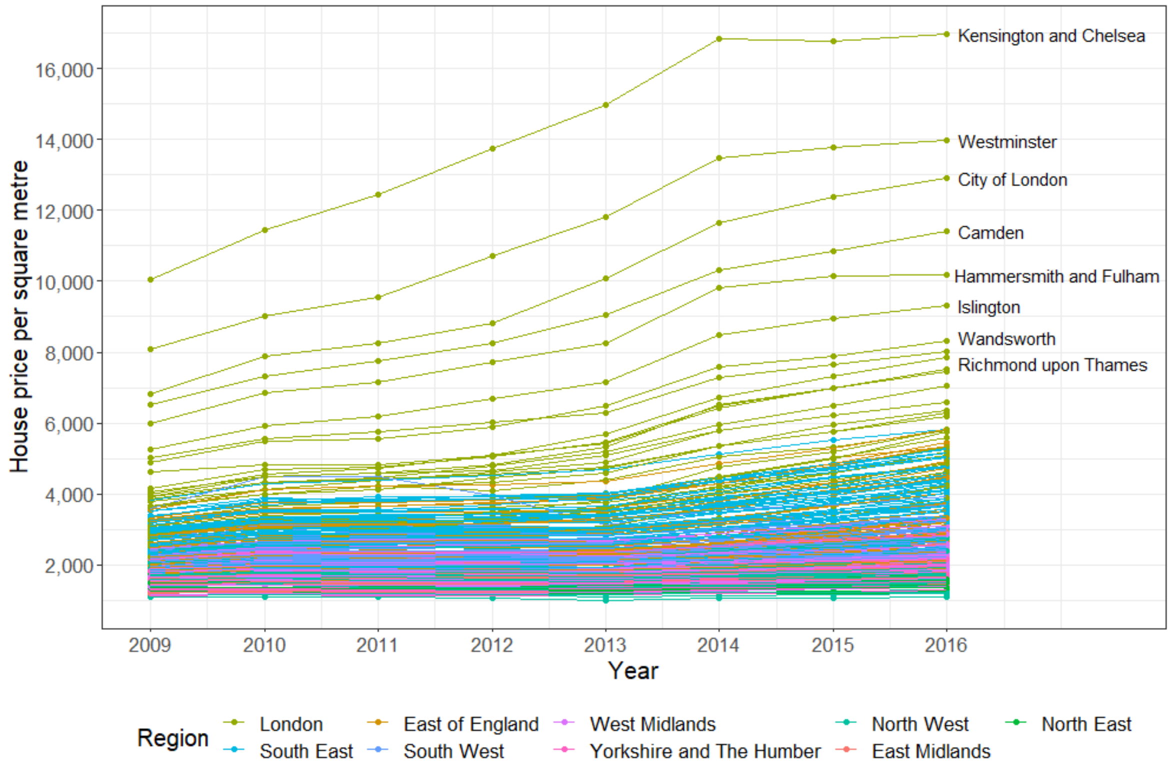


Figure 5.3 House price trends at LA level

Observing the geography of house price differentials at LA level in Figure 5.3 separately by year (Appendix B1), reveals a house price ripple effect starting from inner London and spreading out to peripheral areas. The pattern of house price differentials at LA level is not constrained to the North-South Divide pattern, but a group of contiguous LAs near to national parks in the north also exhibit a relatively high house price.

5.3 Method

This research is divided into three stages with three methods. First, a variance components model is used to explore the space and time effect on house price variance in England between 2009 and 2016, with a greater focus on the time effect in terms of year effect, half-year effect and quarter effect. Second, with a better understand of the time effect of house price variation, GCM is built to offer a model-based description of spatial and temporal patterns of local house price in England between 2009 to 2016. Third, by utilising the estimated increasing rate and starting-price for each LA, from the GCMs, choropleth map and hybrid hierarchical k-means clustering are used to understand the spatial-temporal local housing market pattern in England. A spatial clustered map basing on the hierarchical k-means clustering result is plotted to delineate the spatial-temporal pattern of the LA level house price in England. Details of the three methods are then introduced.

5.3.1 Multilevel Model

5.3.1.1 Variance components model

For geographical research, multilevel modelling is a useful statistical tool to model the relationships which vary in space and over time (Jones, 1991b). In exploring house price variations, variance components model offers a systematic tool to quantify the variances at different spatial scales and time scales. Given England's house price dataset between 2009 and 2016, a three-level variance components model can be built to systematically explore the one spatial effect (i.e. LA) and time effect on house price variance. This is written as:

$$h_{igj} = \beta_0 + l_j + u_{gj} + e_{igj} \quad (5.1)$$

$l_j \sim N(0, \sigma_l^2)$

$u_{gj} \sim N(0, \sigma_u^2)$

$e_{igj} \sim N(0, \sigma_e^2)$

where h_{igj} refers to an individual house price (log scale) for the ith transaction sold at time period g in LA j , β_0 is the fixed term, representing the overall mean house price over the complete time period, and l_j, u_{gj} and e_{igj} are the random term of the variance components model, respectively representing the residuals at the LA level, time period level and individual level. Residuals at each level are assumed to follow normal distributions with zero means and constant variance. l_j measures the extent to which the mean house price in LA j varies from the overall mean house price (β_0), u_{gj} measures the extent to which mean house price at time g in LA j deviates from mean house price in LA j for the whole period. Residuals at the same level are uncorrelated with each other, and residuals at different levels are also uncorrelated with each other.

The three-level variance component (equation 5.1) can extend to four levels to examine the two location effects and one time-effect simultaneously, which is achieved by adding a new random term. House prices in England are quietly similar within the same MSOA for the same year between 2009 and 2016. Given this, a four-level are built to explore the extent of house price variation in the LA, MSOA and time. Equations are shown in equation 5.2:

$$h_{igkj} = \beta_0 + l_j + m_{kj} + u_{gkj} + e_{igkj} \quad (5.2)$$

$l_j \sim N(0, \sigma_l^2)$
$m_{kj} \sim N(0, \sigma_m^2)$
$u_{gkj} \sim N(0, \sigma_u^2)$
$e_{igkj} \sim N(0, \sigma_e^2)$

where h_{igkj} refers to an individual house price (log scale[18]) i happened at time period g in MSOA k and LA j, while β_0, l_j, σ_l^2, σ_u^2 has the same meaning in equation 5.1. m_{kj} is the new added random term and also called MSOA level residual, which measures the extent to which the mean house price of MSOA k deviates from the mean house price in LA j for the whole period, u_{gkj} is still the time residual and now it measures

[18] The reason for using the log scale is to be consistent with the subsequent growth curve modelling.

mean house price difference between mean house price of a given time(e.g. one year) at one MSOA to MSOA's grand mean house price over the whole period. e_{ikj} is the individual residuals, measuring the house price difference between any individual house price to its mean house price within the same MSOA and same time period.

In this four-level variance components model, total house price variance is decomposed into four variance parts ($\sigma_l^2, \sigma_m^2, \sigma_u^2$ and σ_e^2), which assess the variable around the grand mean at the level of LA, MSOA, time and individual property (Jones and Bullen, 1993). σ_l^2 is the variance at LA level, measuring house price differences between LAs over the whole period; σ_m^2 is the MSOA level variance, measuring the price different within-local-authority-between-MSOAs over the whole period. σ_u^2 is the residual variation at time level, which measures the time-to-time (e.g. year-to-year) differences within the same MSOA; σ_e^2 is the individual variance, measuring the house price variability within the same time in the MSOA to which it belongs. Variance partition coefficients (VPC) represents the percentage variance explained by a given level in multilevel model by using the four variance parts (σ_l^2 , σ_m^2 , σ_u^2 and σ_e^2). It ranges from 0 to 1; 0 signifying no between group differences and 1 signifying no within group differences. A higher VPC at a particular level indicates that a greater proportion of total variation is due to differences between the units at that level. The equation for VPC at LA level is presented in equation 5.3, with the following equations for VPC at MSOA level (equation 5.4), time (equation 5.5) and individual level (equation 5.6).

$$VPC_l = \frac{\sigma_l^2}{\sigma_l^2+\sigma_m^2+\sigma_u^2+\sigma_e^2} \qquad (5.3)$$

$$VPC_m = \frac{\sigma_m^2}{\sigma_l^2+\sigma_m^2+\sigma_u^2+\sigma_e^2} \qquad (5.4)$$

$$VPC_u = \frac{\sigma_u^2}{\sigma_l^2+\sigma_m^2+\sigma_u^2+\sigma_e^2} \qquad (5.5)$$

$$VPC_e = \frac{\sigma_e^2}{\sigma_l^2+\sigma_m^2+\sigma_u^2+\sigma_e^2} \qquad (5.6)$$

Three four-level variance components models were used to estimate the extent of the house price (i.e. HPM) variability at LA level, MSOA level and separately three different time scales, each considered by a different model. Level 1 is the individual residential property. Level 2 is the time level but separatly refers to three different time periods (Quarter, half-year and year). Level 3 is MSOA level and level 4 is the LA level. Equations of these three models are listed in Table 5.1. Likelihood ratio test is used to test the significance of the LA effect and time effect in Model 1, 2 and 3. LA effect is first to test between the candidate models in Table 5.1 and its relative two-level variance component model by dropping LA level. Three different time effects are tested through three pairwise likelihood-ratio test between candidate models to their relative two-level variance component model by dropping time level. Meanwhile, Likelihood ratio test is also used to identify which is the best fitted model in Table 5.1

Table 5.1 The candidate four-level variance component models

Model	**Equation**
Model 1	$hp_{iskj} = \beta_0 + l_j + m_{kj} + q_{skj} + e_{iskj}$
Model 2	$hp_{iwkj} = \beta_0 + l_j + m_{kj} + hy_{wkj} + e_{iwkj}$
Model 3	$hp_{idkj} = \beta_0 + l_j + m_{kj} + y_{dkj} + e_{idkj}$
Notes: hp is the log scale of HPM. For example, hp_{iskj} stands for the log of HPM i in quarter period s in MSOA k in LA j. β_0 is overall mean house price across the LAs over the complete time period, l_j is the residuals at LA level, m_{kj} is the residuals at MSOA level, q_{skj} in Model 1 the residual at time level in terms of quarter, hy_{wkj} in in Model 2 is the residual at time level in terms of half-year period, y_{dkj} in Model 3 is the residual at time level in terms of year. e_{iskj}, e_{iwkj} and e_{dikj} stand for individual level residual in each model.	

5.3.1.2 GCM

GCM is a multilevel model using time as a predictor, which fits the trend of repeated-measures data over time vary across different levels (Goldstein, 2010). GCM has been well used in longitudinal study when addressing questions about change (Singer and Willett, 2003; Steele, 2008b; Zaninotto et al., 2009). In house price analysis, house price can be treated as repeated measurement for the same areas (Jones and Bullen, 1993). For example, the HPM of properties (level 1) are recorded from different LAs (level 2). Such a two-level basic GCM can be represented formally using the following equation:

$$hp_{ij} = \beta_0 + \beta_1 t_{ij} + l_j + e_{ij} \quad (5.7)$$

$l_j \sim N(0, \sigma_l^2)$

$e_{ij} \sim N(0, \sigma_e^2)$

where h_{igj} is the individual house price (log scale) for the *i*th transaction in LA j, t_{ij} is the time (i.e. year) of the transaction *i* in LA j. The natural logarithm of the response is used to deal with the technical problems of non-linearity and provides a meaningful interpretation of estimated slope parameter β_1 . β_1 is overall average slope, which is the approximation equal to the overall percentage increases in England over the whole period (2009-2016). β_0 is the overall mean, which is interpreted as the overall house price in England (2009-2016) in terms of a logarithmic scale. The fixed part in the multilevel model is $\beta_0 + \beta_1 t_{ij}$, the random part is $l_j + e_{ij}$. l_j and e_{ij} are the residuals. Residuals at a given level are assumed to follow a normal distribution with zero mean and constant variance. Moreover, residuals at the same level or different levels are assumed to be uncorrelated with one another.

In equation 5.7, all the LAs in level 2 share growth trajectory (β_1). LA growth trends have been observed to vary in in Figure 5.3. GCM can permit this growth varies among LAs by adding a random part $l_{1j}t_{ij}$, the new equation can be written as:

$$h_{ij} = \beta_0 + \beta_1 t_{ij} + l_{0j} + l_{1j} t_{ij} + e_{ij} \quad (5.8)$$

$l_{0j} \sim N(0, \sigma_{j0}^2)$

$l_{1j} \sim N(0, \sigma_{j1}^2)$

$e_{ij} \sim N(0, \sigma_e^2)$

here h_{ij}, β_0, β_1 and e_{ij} have the same meaning as before in equation 5. l_{0j} is the same meaning as l_j in equation 5.7. In the new random term l_{1j} measures the extent to which slope of LA j deviates from the overall slope β_1. The random effects l_{1j} and l_{0j} are assumed to follow normal distributions with zero means, variances σ_{j0}^2 and σ_{j1}^2 respectively, and covariance σ_{j01} . e_{ij} is also assumed to follow a normal distribution with zero mean and constant variance σ_e^2.

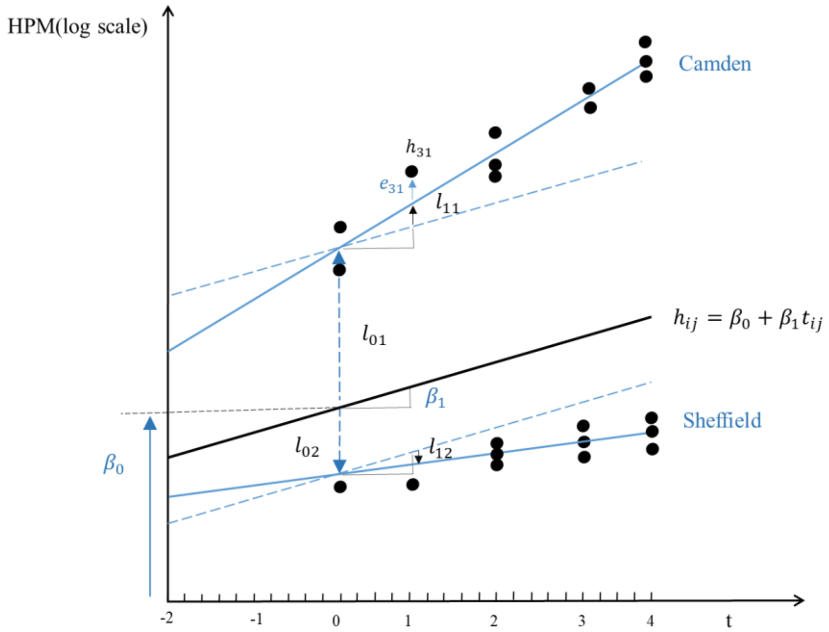


Figure 5.4 A graphical illustration of the two-level GCM (equation 5.8)

Figure 5.4 provides a graphical illustration of equation 5.8 for 22 transactions in two LAs (Camden and Sheffield) in England during five consecutive time periods. Individual TPs are shown as black circles. β_0 is the intercept, which represents grand mean HPM (log scales) in England when time is 0. β_1 represents the overall slope in England across the whole time period, which is an approximation equal to percentage change of the HPM. $\beta_0 + l_{0j}$ measures the intercept for LA j, and $\beta_1 + l_{1j}$ measures house price percentage change for LA j, Camden has a bigger intercept ($\beta_0 + l_{01}$) than the mean house price in England (β_0) with a positive l_{01}, while Sheffield has a smaller intercept ($\beta_0 + l_{02}$) than the mean house price in England with a negative l_{02}. Meanwhile, the slope of the Camden ($\beta_1 + l_{11}$) is steeper than the overall average slope line (the black line) by an amount l_{11}, while Sheffield has a slope ($\beta_1 + l_{12}$) which is smaller by an amount l_{12}. For the house price in Camden and Sheffield, a high intercept is associated with a steep slope. If this pattern holds when all LAs are considered, the intercept-slope covariance will be positive and the group lines (the blue solid lines) will 'fan out'. e_{ij} measures for house price differences for each individual i over the intercept (average LA house price at time 0).

Given Chapter 4 found that HPM within the same MSOA are more similar than HPM within the LA, we need to consider this dependency by treating the MSOA as one random effect. Similar to extension from equation 5.1 to 5.2, equations 5.7 and 5.8 can extend to a three-level GCM by adding in a random term m_{kj}, for details see the Model 4 and 5 in Table 5.2. In Model 4 and Model 5, Level 1 is individual, level 2 is MSOA level and level 3 is LA level. Time variable (t_{ikj}) is centred at the beginning of year 2009 so that the estimated intercept has a more meaningful interpretation (Raudenbush and Bryk, 2002), which is the estimated house price (log scale) in 2009. A likelihood ratio test is used to compare Model 4 and Model 5 to get the better fit model. In this research, the estimated slope for each LA in model 4 or model 5 was named as "estimated HPM percentage change" (LA slope, such as β_1 in model 4 or $\beta_1 + l_{1j}$ in model 5), the estimated intercept is transferred to its natural scale for each LA (exponential $\beta_0 + l_j$ for each LA j) and named as the "starting-price".

Table 5.2 The candidate three-level GCMs

Model	Equation
Model 4	$hp_{ikj} = \beta_0 + \beta_1 t_{ikj} + l_j + m_{kj} + e_{ikj}$ $l_j \sim N(0, \sigma_l^2)$ $m_{kj} \sim N(0, \sigma_m^2)$ $e_{ikj} \sim N(0, \sigma_e^2)$
Model 5	$hp_{ikj} = \beta_0 + \beta_1 t_{ikj} + l_{0j} + m_{kj} + l_{1j} t_{ikj} + e_{ikj}$ $l_{0j} \sim N(0, \sigma_{l0}^2)$ $l_{1j} \sim N(0, \sigma_{l1}^2)$ $m_{kj} \sim N(0, \sigma_m^2)$ $e_{ikj} \sim N(0, \sigma_e^2)$
Notes: hp_{ikj} is the log HPM for transaction i in MSOA k belonging to LA j. t_{ikj} is the time period of the corresponding transaction, time scale is choose according related time scales of the best fitted model among Modes 1 to 3. β_0 is overall mean HPM across all LAs between 2009 and 2016, β_1 is the slope, l_j or l_{0j} is the residual at level 3, m_{kj} is the residual at level 2, e_{ikj} is the residual at level 1. l_{1j} is the random slope at level 3.	

5.3.2 Clustering method

5.3.2.1 Variable standardisation and test the clustering tendency

Utilising the better fit GCMs in the previous section, each LA in England has its estimated HPM percentage change and starting-price. Initially, these two attributes were standardised to weighted these two attributes are equally considered during the clustering process (Dennett, 2010; Everitt et al., 2011). Estimated HPM percentage change and starting-price are roughly normally distributed data, *z*-score is chosen to standardise the two attributes. The *z*-score standardization formula is defined as:

$$Z_i = \frac{x_i - \mu}{\sigma} \qquad (5.9)$$

where x_i is the original data value, and μ and μ are the sample mean and standard deviation, respectively.

5.3.2.2 Hybrid hierarchical k-means clustering method

K-means is one of the most popular clustering methods developed over 60 years (Jain, 2010). It is an unsupervised Machine learning algorithm for partitioning a dataset into *k* clusters; clusters are defined with a high intra-class similarity and low inter-class similarity. The working process of k-means clustering is fast and easy to understand. It starts by randomly selecting *k* observations from the dataset as the initial centres for the clusters and *k* is the cluster group determined in advanced. Then each of the rest of the observations is assigned to its closest centroid. After this assignment, the algorithm computes the mean value of each cluster as the new cluster centres. All the objects are reassigned again using the updated centres. The cluster assignment and centroid update steps are iteratively repeated until the current clusters formed are the same as those obtained in the previous iteration (Kassambara, 2017).

However, different initial cluster centres can lead to different clustering results because k-means clustering is very sensitive to this initial random selection of cluster centres (Everitt et al., 2011). Defining initial points for *k*-means is one approach to overcome this *k*-means disadvantage (Khan and Ahmad, 2004). Hybrid hierarchical k-means clustering method follows this idea by setting the cluster centres from the hierarchical clustering as the initial cluster centres to improving k-means result (Kassambara, 2017). Therefore, Hybrid hierarchical k-means clustering method are used to sort the trends of an LA's HPM by considering the both HPM percentage change and starting-price.

5.3.2.3 Determining the number of clusters

Deciding the number of clusters is important but quite challenging, because different methods will result in a different optimal number of clusters (Kassambara, 2017), this has become increasing clear that seeking the best clustering might indeed be futile as there is no definitive answer (Jain, 2010). The basic idea of K-means algorithm defined clusters is to minimize the total within-cluster variances, thus understanding the extension of the total within-cluster variation change along with different increasing the number of clusters can offer a guidance to how different clustering performs. The Elbow method follows this idea by plotting total within-cluster variation against the number of clusters k and consider the number of clusters when adding more one cluster does not decrease so much of the total within-cluster variation (Han, 2011; Kassambara, 2017). In this research, the clustering method is used to assist an exploration of the similarity of the LAs' house price between 2009 and 2016 and classify 326 LAs into small groups based on two estimations (starting-price and estimated HPM percentage change). The more clusters group, the less total within-cluster variation and the clearer the picture of LA house price difference becomes; thus the number of clusters is selected at the point before which total within-cluster variation decreases almost equally when clustering one more group.

5.4 Results and discussion

Models 1 to 5 were run in MLwiN 3.03 (Charlton et al., 2019) using the Iterative Generalized Least Squares (IGLS) algorithm. The likelihood ratio test on LA, MSOA and time random effects for each of the Models 1 to 3 are associated with effectively zero p-values, revealing that LA, MSOA and time variance are separately significant in these three models. Similarly, the Likelihood ratio test on LA, MSOA effect in Model 4 and 5 also results in a separately effectively zero p-value. Meanwhile, Model 3 with the lowest deviance among the Models 1 to 3 reflects the best fit model in the four-level variance models. The year was therefore chosen as the time scale in Models 4 and 5. A likelihood ratio test reveals that Model 5 is preferred over Model 4 (LR = 175386, $p < 0.001$). All the results discussed below are based on the estimated values from the above five multilevel models. Clustering sections are conducted in R. Choropleth maps are plotted in ArcGIS 10.6.

5.4.1 LA and time effects on house price variation in England (2009-2016)

Results of four-level variance component models are listed in Appendix B2. Table 5.3 shows the VPC result of these three models, with VPC at the same level for all three models are the exactly the same when rounding to 2 decimal places. The VPC at time level is the same for three different time periods, which is 0.05. This means that there is no different for time effect in terms of three different time periods (i.e. quarter, half-year and year) in England house price variance. Comparing to the LA and MSOA level effect on the total house price variance, time effect is quite tiny (only accounting for 5% total variance). Due to the tiny VPC at time level, time are treated in the fixed effect not random effect anymore in the following analysis. Moreover, deviance of the Model 3 is smallest and this indicated that a one-year time scale fits the data better than the other two time categories. Therefore, subsequent analysis exclusively uses a one-year time scale.

Table 5.3 VPC statistic for Model 1, Model 2 and Model 3

Model 1		**Model 2**		**Model 3**	
Level	**VPC**	**Level**	**VPC**	**Level**	**VPC**
LA level	0.59	LA level	0.59	LA level	0.59
MSOA level	0.12	MSOA level	0.12	MSOA level	0.12
Quarter level	0.05	Half-year level	0.05	Year level	0.05
Individual level	0.24	Individual level	0.24	Individual level	0.24
Deviance	1,428,443	Deviance	1,338,665	Deviance	1,287,883

In Model 3, VPC at LA level is biggest (0.59); this indicates that 59% total HPM variance (log scale) between 2009 and 2016, in terms of log scale lies between LAs; in other words, house price difference among LAs in England is large. Meanwhile, 12% of total HPM variance lies between MSOAs within the same LA. Of the remaining 29% of variance, only 5% is due to year difference. 25% of total HPM variance occurs at the individual level, which could be due to differences between individual properties (e.g. plot size, property quality), after controlling for total floor size effect.

5.4.2 LA house price and average change

Table 5.4 summaries the model results from Model 4 and 5. Due to a large deviance decrease between Model 5 and 4, the Likelihood ratio test gives a near zero p-value. This suggests that Model 5 fits the data significantly better than Model 4. This also reveals that an LAs' house prices growth trends do vary at LA level.

Table 5.4 Model result of GCM[19]

Parameter	Model 4		Model 5	
	Estimate	S.E.	Estimate	S.E.
β_0 Intercept	7.5613	0.0237	7.5639	0.0199
β_1 (Year-2009)	0.0386	0.0001	0.0379	0.0013
σ_{l0}^2 between LA variance	0.1806	0.0144	0.1262	0.0102
σ_{l01} Intercept-slope covariance	-	-	0.0061	0.0006
σ_{l1}^2 Slope variance	-	-	0.0006	0.0000
σ_m^2 between MSOA variance	0.0369	0.0007	0.0373	0.0007
σ_e^2 Individual variance	0.0789	0.0001	0.076	0.000
Deviance	1,438,463		1,263,077	

In Model 5, the covariance between the intercept and slope is 0.0061, suggesting a

[19] Model 4 fits better than its relative three-level variance components model (deviance is 1964419) according to the Likelihood ratio test. Besides, we did not continue to set the slopes vary at MSOA level due to a super tiny slope variance will being observed at MSOA level (0.0001). This reveals that the house price growth trend is quite similar within the same LA.

positive relationship between the LA slope and intercept. In other words, expensive LA is growing relative faster than cheap LA. Meanwhile, as the slope variance at LA is also positive (0.0006), this reveals a 'fanning out' growth trend (Figure 5.5) exists in the local housing market (housing market at LA level) in England between 2009 and 2016. Moreover, intercept variance (σ_{l0}^2) at LA level is bigger than the slope variance. This advance reveals a huge difference between the house price among LAs in 2009.

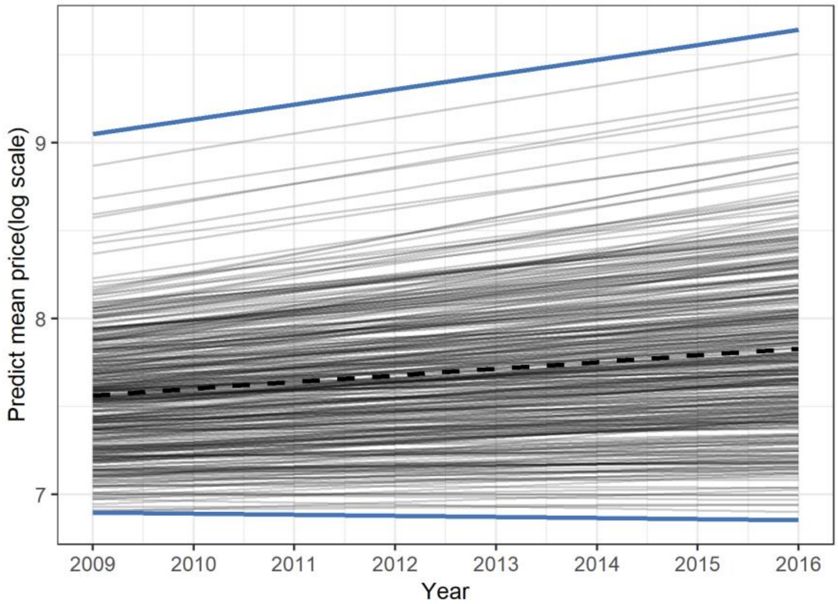


Figure 5.5 England's fanning out growth trend at LA level

Figure 5.5 shows the estimated growth curves for each LA in Model 5. Each line stands for one LA. To further explore this growth trend, Figure 5.6 is created from Figure 5.5 by plotting the intercept and slope for each line. The intercept has been transformed back to its natural scale for each LA, and thus refers to the starting-price. Each point stands for one LA and is coloured by region. The black dashed lines indicate the England's starting-price (1927 £/m^2) and its estimated HPM percentage change (3.79%.). It is clear that the fanning out is not simple, as HPMs in expensive LAs grew relatively more quickly than in cheaper LAs, between 2009 and 2016.The majority of

LAs in England show an increasing trend between 2009 and 2016; only 13 LAs, in the North East, the North West and Yorkshire and the Humber, show a small decreasing trend over the same period. They are Hartlepool, Middlesbrough, Redcar and Cleveland, County Durham, Sunderland, Blackburn with Darwen, Blackpool, Allerdale, Carlisle, Eden, Burnley, Scarborough and Bradford. The top eight most expensive LAs (Kensington and Chelsea, Westminster, Camden, City of London, Hammersmith and Fulham, Islington, Richmond upon Thames and Wandsworth), having HPM over 4000 £/m^2 in 2009, show a greater than 8% price increase in the following 7 years. But they did not display the highest HPM percentage increase. The City of London displays the highest HPM percentage increase in this cluster but this ranks only fifth among the LAs in England. The top 4 highest percentage increase LAs (Waltham Forest, Hackney, Lewisham and Lambeth) exhibit a higher than 10% HPM increase.

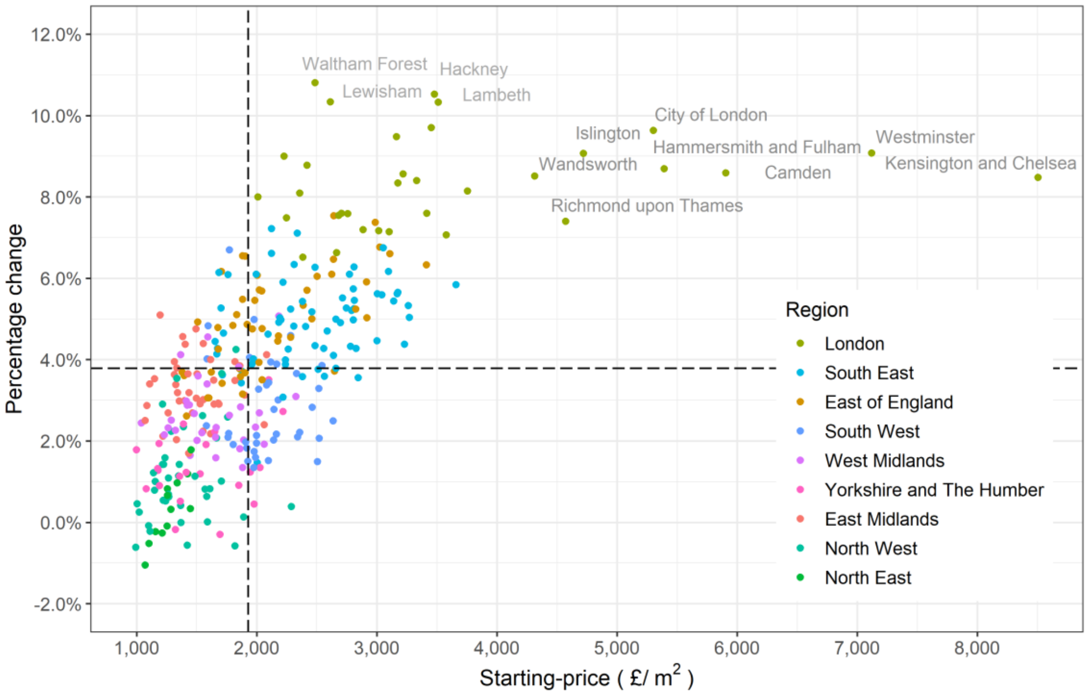


Figure 5.6 The relationship between starting-price and HPM percentage change based on Model 5

The overall house price percentage increase in England between 2009 and 2016 is 3.79%. To better understand the difference of HPM percentage increase at LA level, Figure 5.7 separately plots Figure 5.6 by region. London has a higher percentage increase than the overall level for England (the horizontal dashed line). All London's LAs have more than 6% increase between 2009 and 2016. LAs in the East of England and South East show a moderate HPM percentage between 2% and 8%. Moreover, these LAs are quite diverse in terms of the house price percentage increase, but the majority of them are over England's increase level. LAs in the East Midlands, South West and West Midlands show a small HPM percentage increase around the England's average level (the horizontal dashed line) with a percentage increase between 2% and 6%. With the exception of the Trafford, the other LAs in the North West and Yorkshire and The Humber have a small percentage increase below the England level (3.79%.). LAs in the North East show quite small changes in HPM, which is generally below 2% and around zero.

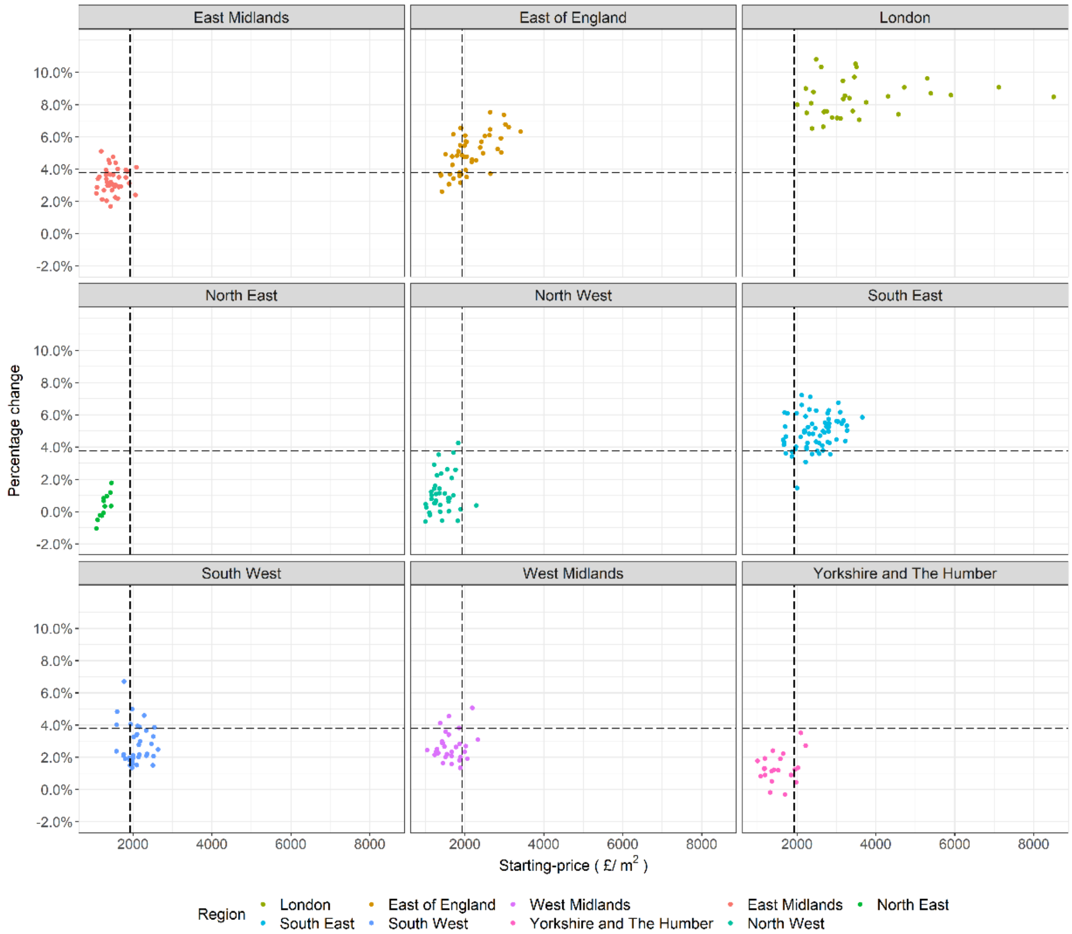


Figure 5.7 The relationship between starting-price and HPM percentage in different regions

The LAs with a starting-price and HPM percentage increases above the England average (top-right quadrant of Figure 5.7) show quite diverse behaviour compared to all other LAs. Within this group nearly all are within London, the South East and the East of England. These two dimensions in house price trend are separately plotted as y axis by region in Figure 5.8. The x axis in Figure 5.8 is the LA's rank order based on the y value and the dashed lines represents the England wide level of the y value. Obviously, the big differences of LAs within London are the core contribution for the LAs' variation in starting-price and percentage change. Looking at the LAs' HPM percentage increase, London's LAs exhibited increases of more than 6% between 2009 and 2016, which is far greater than the England level (3.79%). The majority of LAs in the East of England and the South East exhibit moderate increases of from 3.79% to 7.6%. Only the Isle of Wight shows relatively small price increases (1.47%) compared with the rest of the LAs in South East. LAs in the South West, West Midlands and East Midlands saw small increases at around the average level for England, between 2% and 6%. With the exception of Trafford, the remaining LAs in the North West and Yorkshire and The Humber saw small percentage increases, below England's average. LAs in the North East saw only very small HPM changes, generally below 2%, with fewer LAs showing a decreasing overall price change. Meanwhile, the LAs' starting-price pattern within the same region shows a slightly different pattern as LAs' HMP percentage increases. For example, the Isle of Wight shows a similar starting-price to the rest of LAs in the South East, but it has a relatively small percentage increase. LAs in the East Midlands generally have starting-prices below the England level, but the HMP percentage change in some LAs are above the England level.

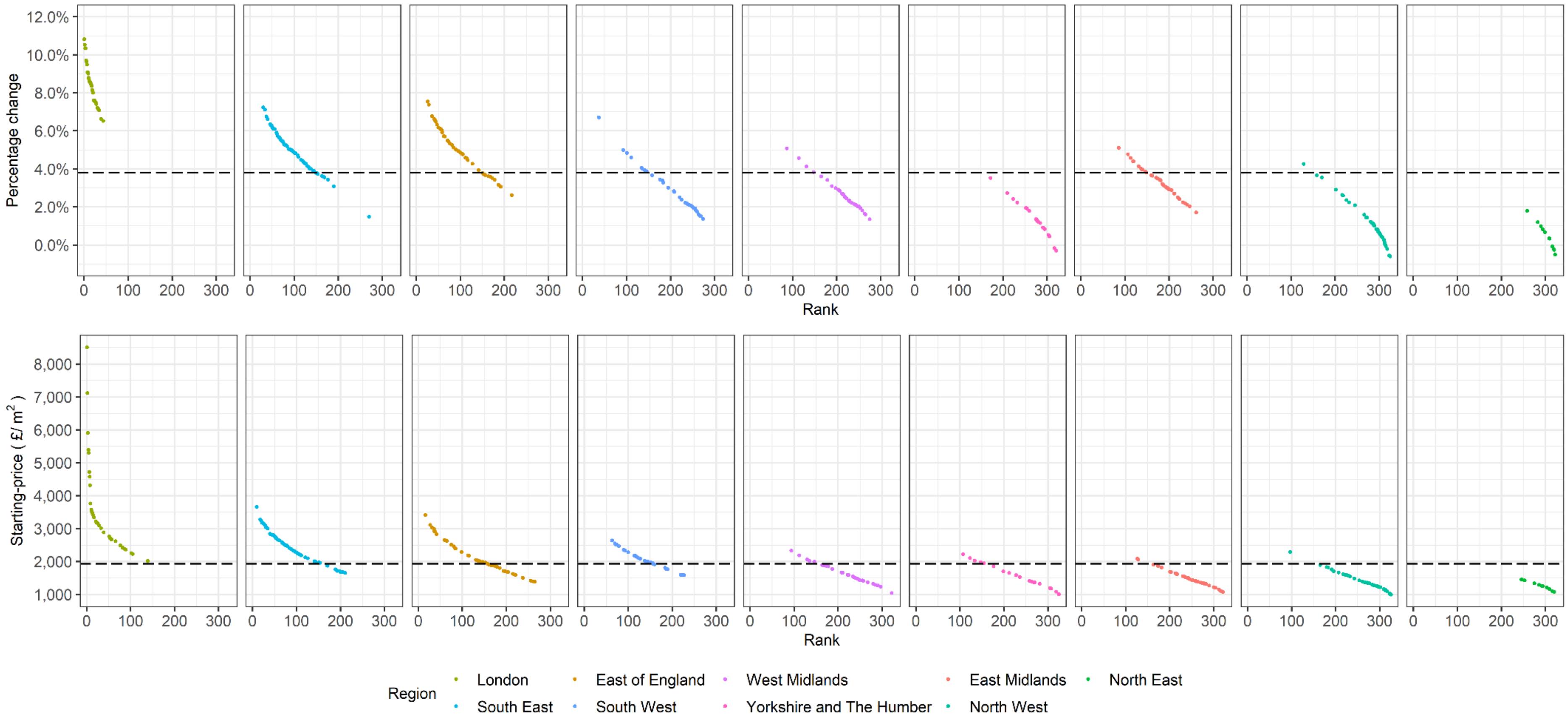


Figure 5.8 LAs' starting-price and percentage change in England by region.

5.4.3 Spatial pattern difference in LA's starting-price and percentage increase in England

Figure 5.9 represents the spatial pattern of average HPM percentage increase in England over the study period. LA HPM percentage changes are crudely sorted into 6 classifications with the same percentage change value interval in Figure 5.6. There are two obvious gradient (ripple) patterns of percentage change at LA level. One is centred on London and the other is centred on Bristol.

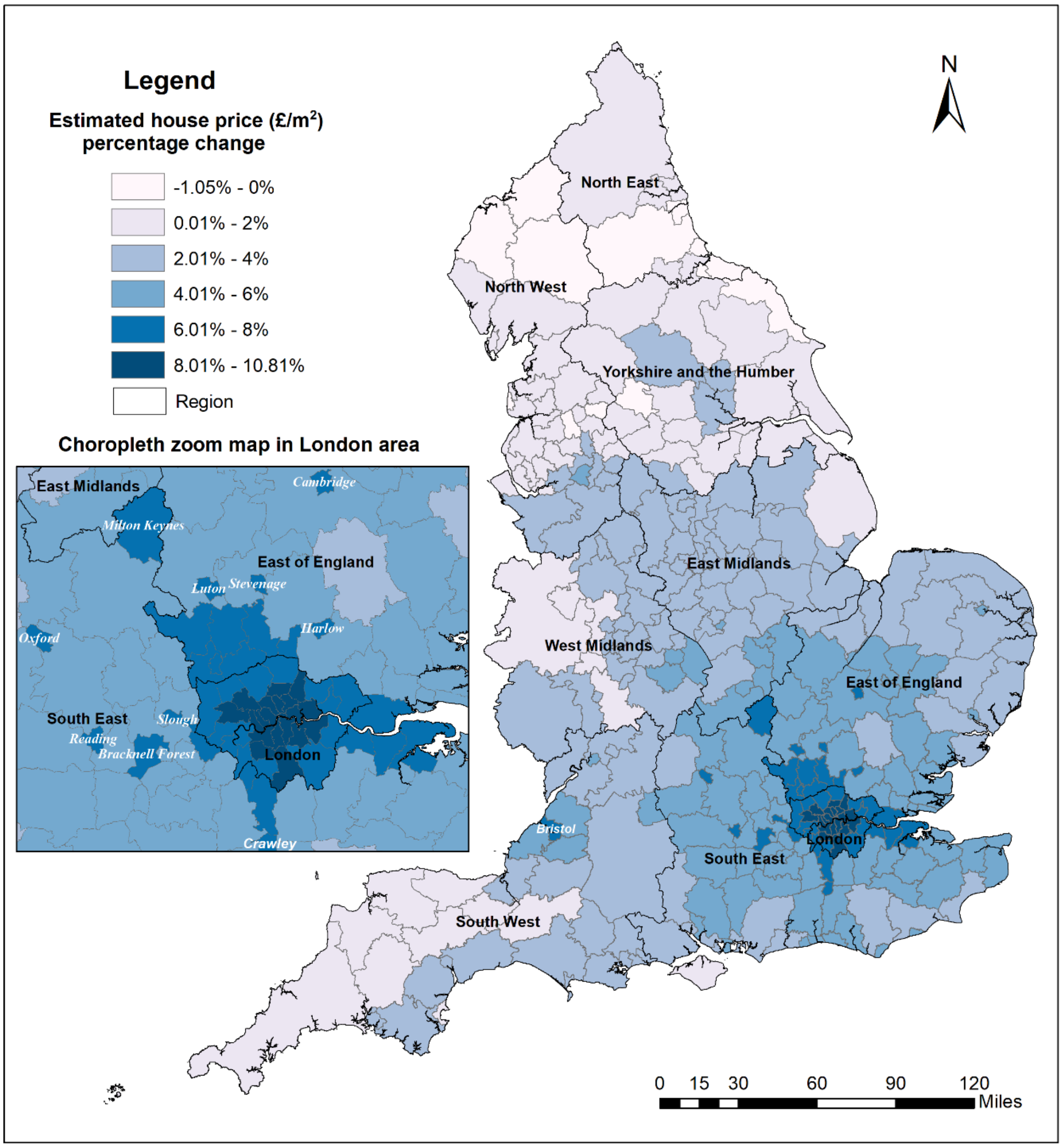


Figure 5.9 The spatial pattern of LA's house price percentage change

In London and its nearby housing market, house price percentage change follows a gradient pattern with high increasing percentage at the centre of London, decreasing in price away from the centre. This decreasing trend generally correlates with distance from centre London ,with an exception of nine LAs (labelled on the inset map in Figure 5.9). These nine LAs show a higher percentage increase (over 6%) compared to neighbouring LAs, and their travel time to London is around an hour. The underlying reasons that the housing markets of these nine LAs differ from their neighbouring areas are likely to vary from case to case. One potential reason for the high percentage house price increases in Milton Keynes, Luton, Stevenage and Harlow could be their role as London commuter towns; these areas have a high proportion of people who work in London (Figure 5.10). Figure 5.10 represent the map of percentage of outside travel to work in London against the total outside travel to work at LA scale; based on Census 2011 data. This map is aggregated travel to work data (Table WU03EW) in the Census 2011, at each LA unit, and then treated all the LAs in London as one unit. The proportion of extra-LA commuting that goes to London refers to the number of people commuting outside of their home LA to work in London divided by the number of people commuting outside of their home LA to work. Meanwhile, the figure also represents the geography of Green Belt in England. (Mace et al., 2016; Smith, 2017) within relatively easy commuting reach of London. The reasons for the higher percentage increases in Oxford and Cambridge could be due to local green belt planning constraints or their status as prestigious university citeis (Mace et al., 2016; Smith, 2017) within relatively easy commuting reach of London. Higher percentage HPM increases in Reading and Bracknell Forest may be due to their technology industries and the fact that both are well-connected to London by both the M3 and M4 motorways, as well as fast rail links (Hodson, 2019; Holland, 2019; Osborne, 2016). Indeed alongside Crawley in Sussex which also displays higher percentage HPM increases, many of these residential areas were developed in the post-war wave of new town building designed to re-house London families and have always retained an association with London through these displaced populations and commuting links.

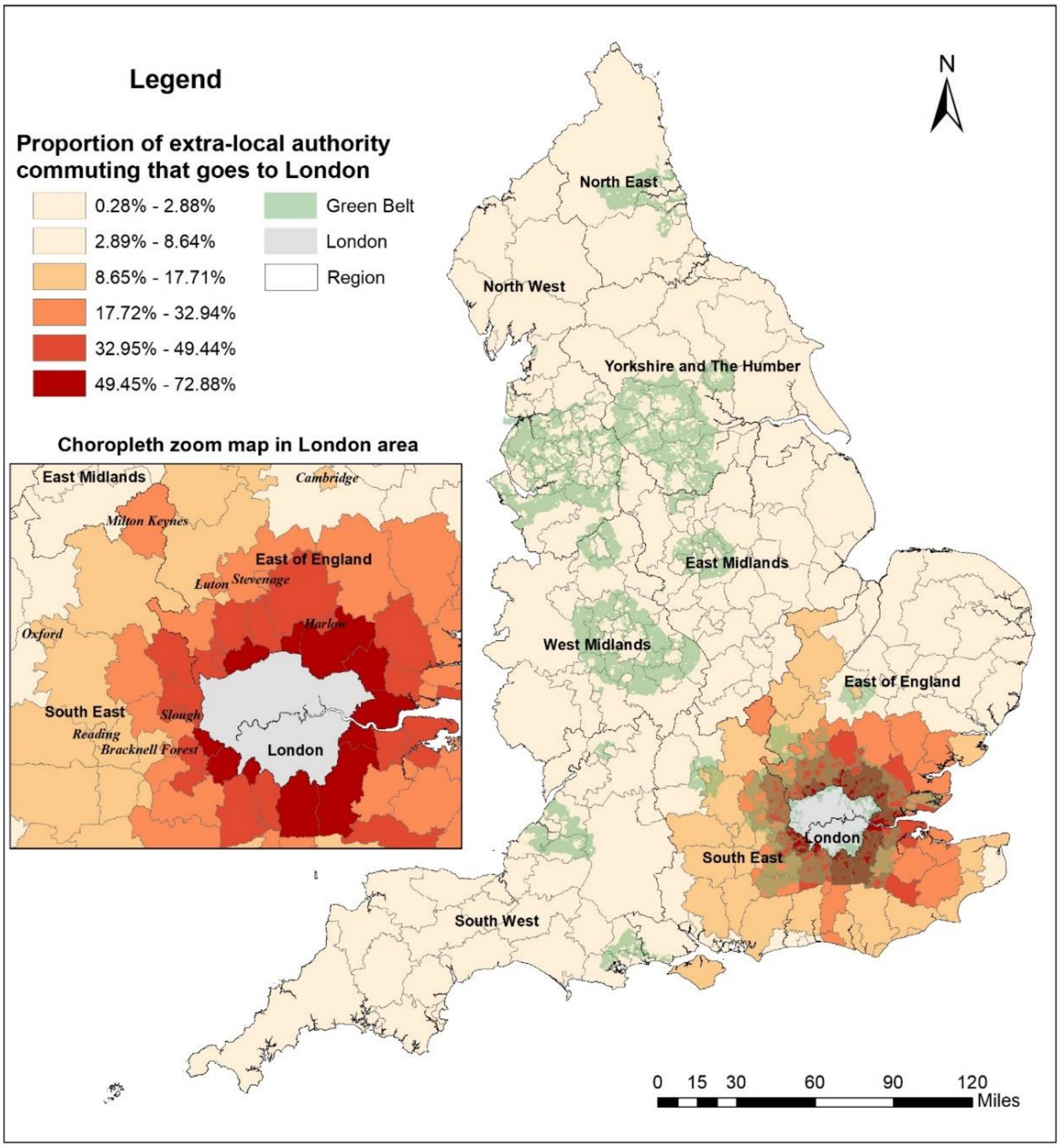


Figure 5.10 Percentage of outside travel to work at London against the total outside travel to work

The LA's HPM percentage change in and around Bristol exhibits another gradient radial pattern, with a high increase in Bristol and a decreasing percentage change as one moves away from the centre, as seen in Figure 5.9. Bristol is a tech hub for the electronics, creative media and aerospace industries (Card, 2014; Ismail, 2018). The pattern observed around Bristol may relate to commuting to work patterns, in the same way that the London effect appears to (Rae, 2017). Bristol may also be influenced by London as it is possible to commute to London within 75 mins (Chi, Dennett, Morphet, et al., 2020). Although these areas have high house price percentage increases, their starting-prices were not as high as those in London and its nearby housing market, as shown in Figure 5.11.

Figure 5.11 demonstrates the spatial pattern of the starting-price at LA level, corresponding to the 1,000 £/m^2 interval in Figure 5.6. There are only two LAs showing starting-price between 6,000 £/m^2 but lower than 8,000 £/m^2 and these two LAs were not sub-classified further. 89% of LAs in England have starting-prices between 1,000 £/m^2 and 3,000 £/m^2 level, with 37% of them being over the 2,000 £/m^2 level. Thirty-five of the remaining LAs, representing almost 11%, have starting-prices over 3,000 £/m^2. These 35 LAs are all located in or near London.

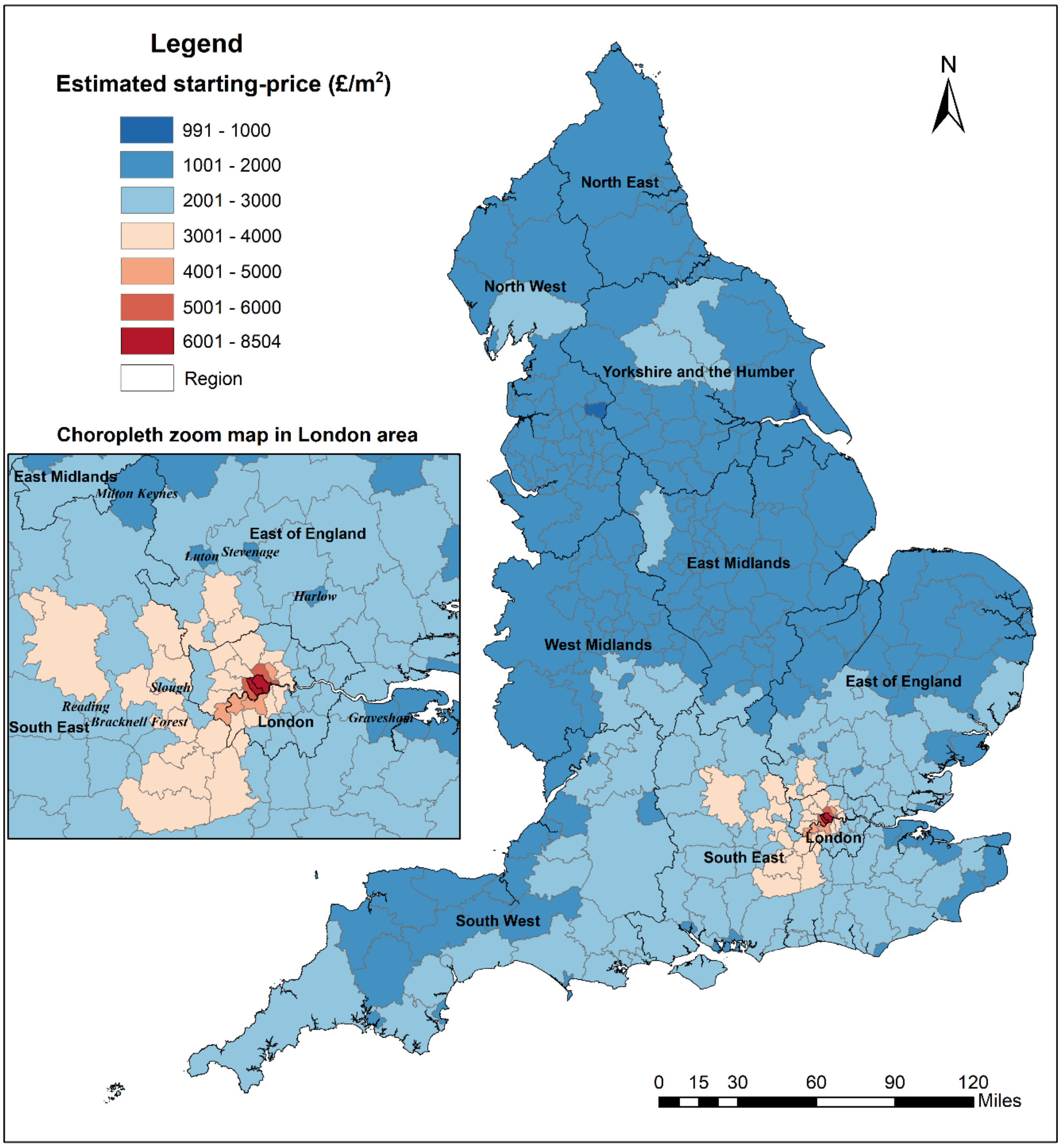


Figure 5.11 The spatial patterns of LA starting-price [20]

Looking at the geography of the starting-price at LA level, HPMs display more complex patterns than would be suggested by the simplistic notion of a "North-South divide". In the south of England, fourteen LAs on the southeast coastline and southwest coastline have HPMs under 2,000 £/m^2, which is relative cheaper compared to their neighbouring LAs. These are Dover, Eastbourne, Gravesham, Hastings,

[20] The log scale of estimated house price in 2009 (intercept) for each LA is transferred to its nature scale and named as starting-price of each LA in England.

Shepway, Medway, Swale, Thanet, Southampton, Gosport, Portsmouth, Weymouth and Portland, Havant and Torbay. Conversely, in the North of England, five LAs display higher HPM s than their neighbours, with HPMs over 2,000 £/m^2: Derbyshire Dales in the East Midlands, South Lakeland in the North West, and Hambleton, Harrogate and York in Yorkshire and The Humber. Burnley in the North West and the City of Kingston upon Hull in Yorkshire and The Humber exhibit house prices below 1,000 £/m^2. The estimated mean HPMs of all other LAs in the North of England lie between 1,000 £/m^2 and 2,000 £/m^2.

Comparing the spatial pattern of the house price percentage increase map (Figure 5.9) and the starting-price (Figure 5.11), Luton, Stevenage and Harlow represent relatively higher percentage house price increase, but a relatively lower estimated HPM in 2009 compared to those neighbourhoods. Similarly, LAs near Bristol show high HPM percentage increases, but their starting-prices are not as high as those in London and its nearby housing market.

5.4.4 The spatial-temporal cluster pattern LA house price in England

Figure 5.12 displays the total within-cluster variation change when adding one more group by using the Hybrid hierarchical *k*-means clustering method. It seems clear that the decrease is close to uniform after group 5 and the decrease is small after group 10. It is not necessary to cluster into more than 10 groups.

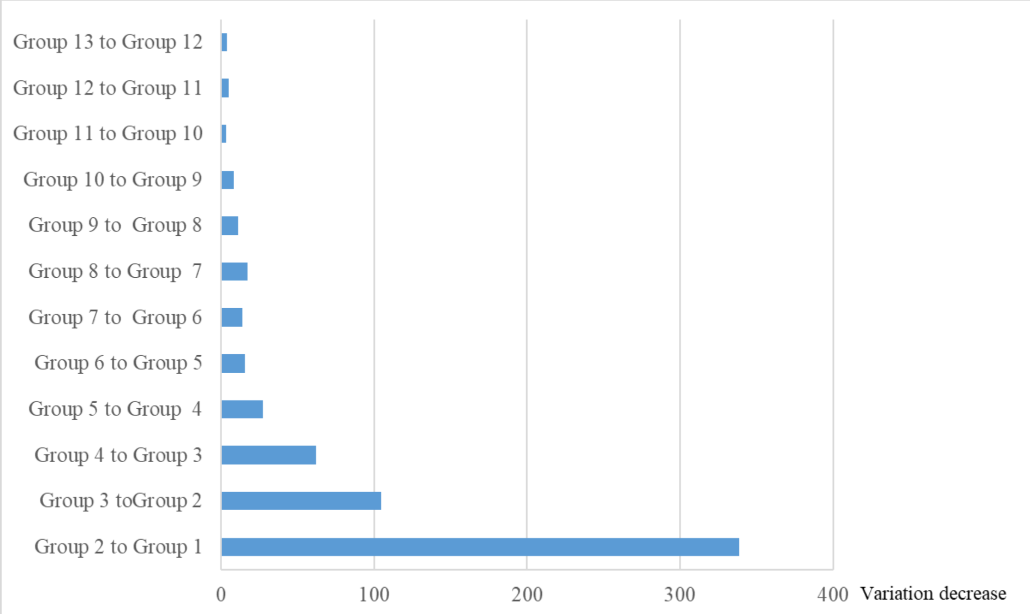


Figure 5.12 Total within-cluster variation decreases after adding one more group

Since only two variables are considered in clustering, it is easy to visualize all the hybrid hierarchical k-means clustering results smaller than 10 groups to choose the cluster number in this research. Figure 5.13 represents the hybrid hierarchical k-means clustering results for the number of clusters which are lower than 10. Each point represents one LA and is coloured the same within the same group. Looking at cluster results for a different number of clusters offers a better understanding of cluster and assists in choosing the final number of clusters. Moving from 2 clusters to 3 clusters, the group 2 in 2 clusters are divided into two new groups (i.e. group 2 and group 3 in 3 clusters) with a few LAs in group 1 assigned to the new group 2. The new added group 3 represents the group of LAs with a high house price in 2009 and a high percentage increase between 2009 and 2016. While moving from 3 clusters to 4 clusters, LAs sorted as group 3 in 2 clusters are exactly the same as group 4 in 3 groups and LAs in the old group 2 and 1 are divided into 3 new groups. Moving from 4 clusters to 5 clusters follows a similar pattern as moving from 3 clusters to 4 clusters, group 1 to 3 in 4 clusters are divided into four new groups (new group 1 to 4) and two LAs in the old group 4 are assigned into the new group 4. The cluster pattern changed a lot from 4 clusters to 5 clusters but afterwards the change is not large. For example, the change from 5 cluster to 6 clusters are the majority which happened in the old group 2 and 3. These two groups continued to be divided into three new groups. The old forth group are general keeps same with a bit points are assigned to the other group. All in all, 5 clusters was chosen as the final cluster result in this research, as the cluster pattern did not change so much when adding one more cluster.

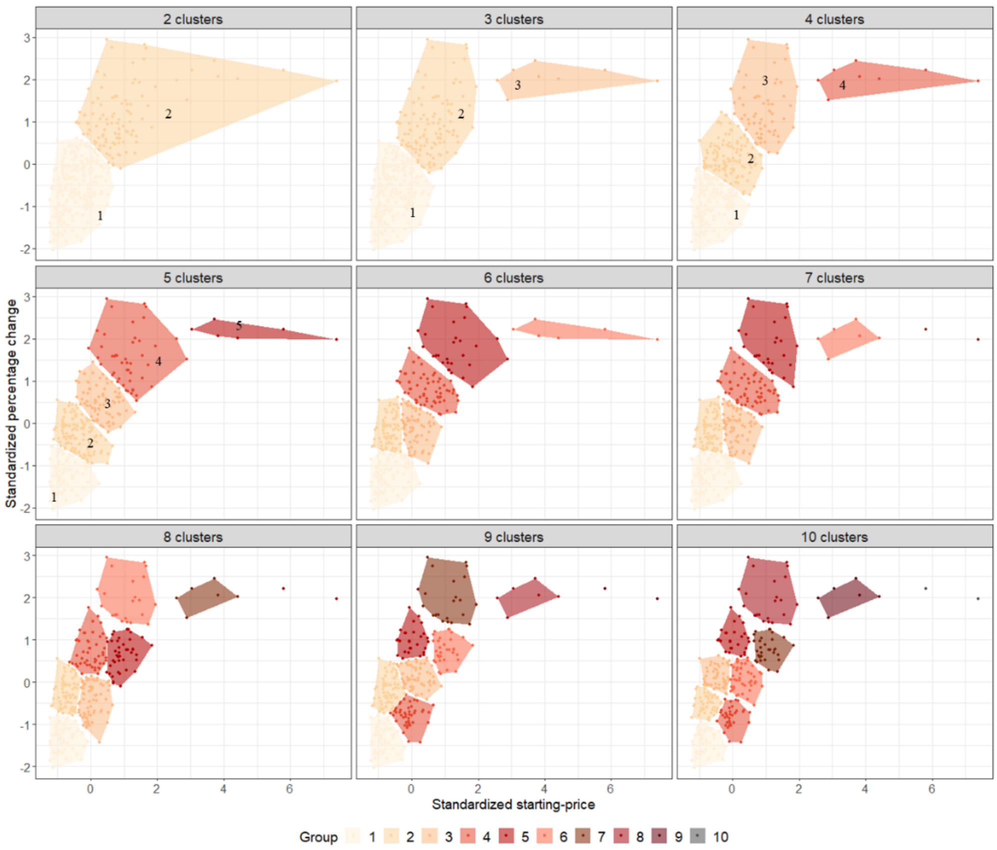


Figure 5.13 Hybrid hierarchical k-means clustering results for clusters number below 10

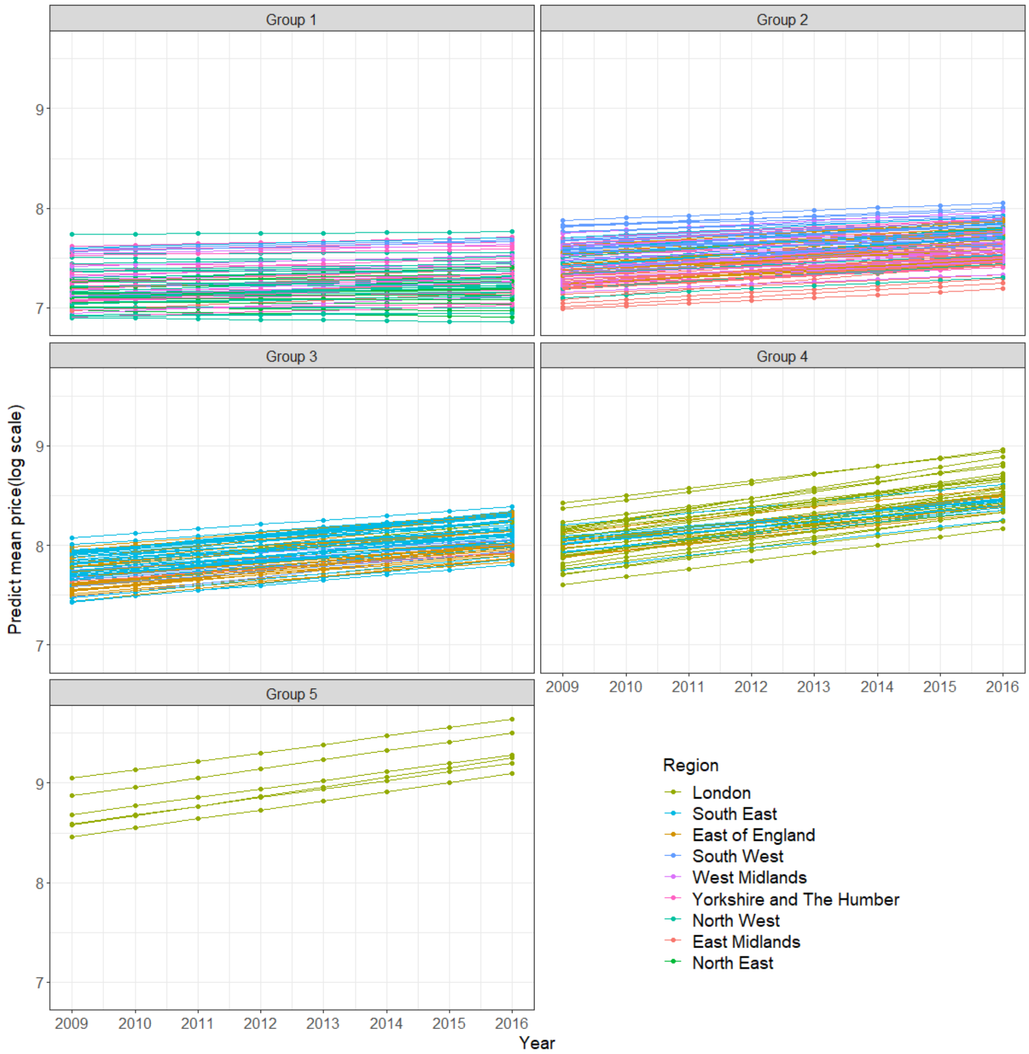


Figure 5.14 Clusters result of house price growth trend at LA level

Given the five clusters result in Figure 5.13, 326 LAs in England are sorted into 5 groups according to the difference in starting-price in 2009 and the average percentage increase in the next seven years. Figure 5.14 separately plots Figure 5.5 according to the five clusters result to better understand the LAs' house price trend in England in each group. Each solid line represents one LA and is coloured by region. The LAs sorted into group 5 are located in London. They are Kensington and Chelsea, Westminster, Camden, Hammersmith and Fulham, City of London and Islington. These areas are defined as an expensive house price area with high increase. The LAs sorted into group 4 are located in three regions (South East, East of England and

London), which have a relatively low starting-price compared with group 5 with a house price increase over 5%. Therefore, these groups are defined as areas of very expensive house price with high increase. Ten LAs in group 4 increase more than LAs in groups 5. They are Waltham Forest, Hackney, Lewisham, Lambeth, Southwark, Haringey, Newham, Greenwich, Ealing and Wandsworth. LAs in group 4 represent a relatively small house price increase and lower starting-price compared to group 5. From this, we define these areas as moderate house price areas with moderate increase. LAs in group 2 have a relatively lower starting-price and house price increase than group 3. Due to the starting-price in group 2 ranging from 1,083 £/m^2 to 2,636 £/m^2 and the percentage increase being lower than 5%, group 2 is defined as the moderate house price area with small increase. LAs in group 1 have the lowest starting-price and small house price change, thus these areas are named as cheap house price with very little change area. Table 5.5 summaries the characteristics of each cluster group and details of the LA, along with its cluster group are listed in Appendix B3.

Table 5.5 A summary of the LA house price cluster

Group	**Number of LAs**	**Type**	**Range of the house price percentage change**		**Range of starting-price**	
			min	**max**	**min**	**max**
1	75	Cheap house price area with very little change	-1.05%	2.50%	991	2287
2	121	Moderate house price area with small increase	1.50%	5.10%	1083	2636
3	78	Moderate house price area with moderate increase	3.29%	7.22%	1685	3229
4	46	High house price area with high increase	5.04%	10.81%	2010	4570
5	6	Very expensive house price area with high increase	8.48%	9.63%	4717	8504

Figure 5.15 represents the geography of house price spatial-temporal pattern by

mapping the 5 clusters result. The area coloured with a darker shade indicates a LAs with a relatively high starting-price and high overall HPM increase between 2009 and 2016. The HPM spatial-temporal pattern in England presents a gradient pattern with expensive house price and high increase starting in inner London but decreasing as distance from the centre increases. This is a kind of London ripple effect current research has observed at regional level. London's high HPM influences its surrounding areas and this influence decreases as commuting distance increases. What is more, this influence is not equally the same in different directions, with a stronger influence in the western parts of the London and the LAs located to the west of London, than in LAs located to the east.

Observing the choropleth zoom map in London in Figure 5.15, the six LAs in group 1 are all located in inner London. The 46 LAs in group 2 occupy 14% of LAs in England. These LAs are mainly located in London and particularly London's western contiguous areas. Oxford and Cambridge are not close to the other LAs in group 2 but show a more similar house price trajectory in terms of the starting-price and average house price increase between 2009 and 2016. As the house price trajectory in Havering is closer to the trajectories in group 3, only Havering in London is sorted into group 3. LAs in group 3 are surrounding the LAs in group 2 and spread out. Moreover, they are all under 2 hours commuting time to London. A couple of LAs in or near the City of Bristol are sorted into group 3. These areas are not only influenced by London but also influenced by Cardiff in Wales (Bowlinson, 2019). Following London's ripple effect, LAs surrounding the group 3 area display a relatively small starting price and overall house price increase; these are sorted into group 4. There are 121 LAs in group 2, accounting for one third of the LAs in England. LAs sorted into group 1 are mainly located in the north of England. These areas have cheaper house prices with very litter house price changes, comprising 23% of the LAs in England.

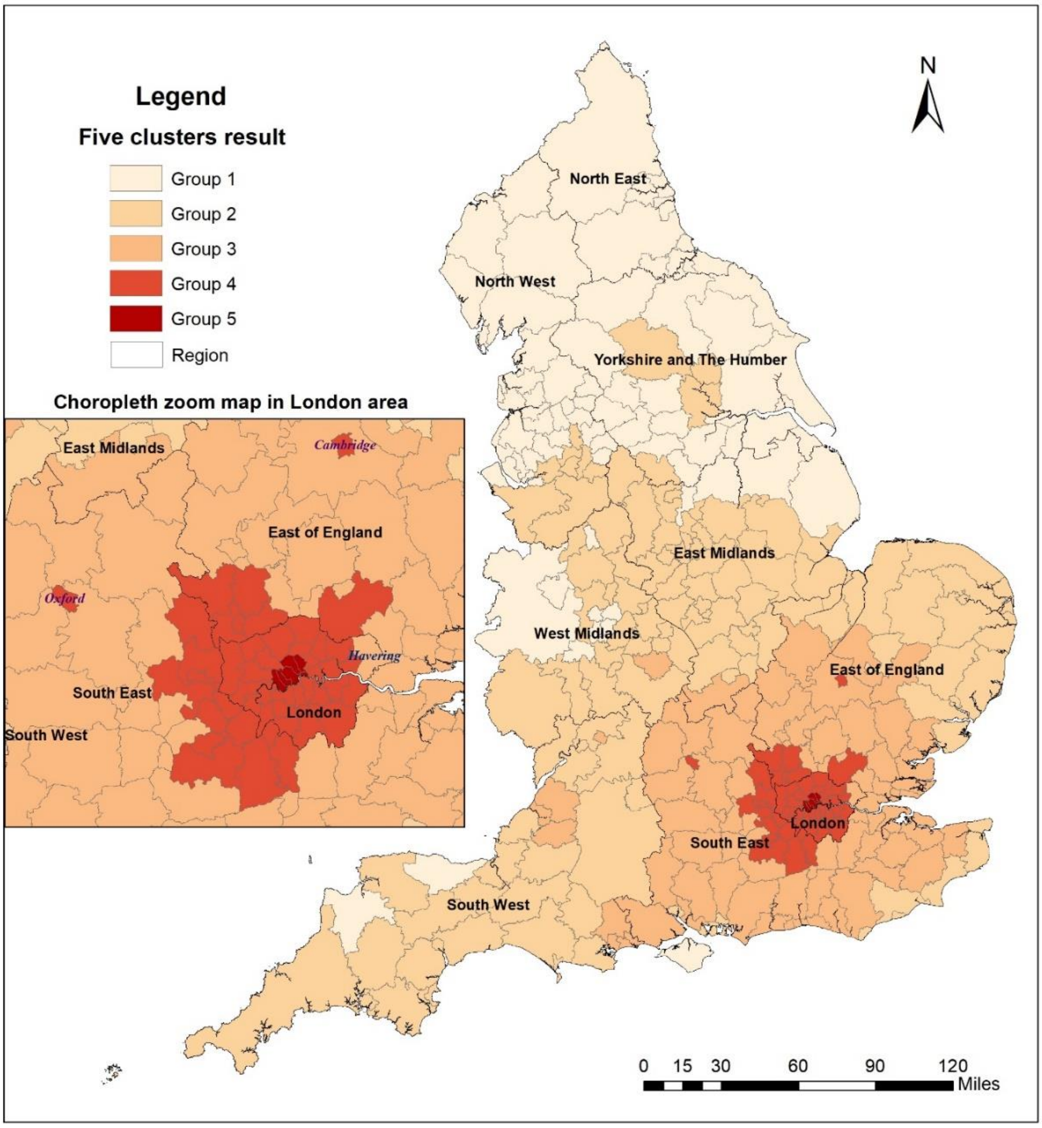


Figure 5.15 Five clusters result of LA spatial-temporal house price in England

5.5 Conclusions

This research takes a first step to systematically explore the spatio-temporal pattern of house prices at LAs level in England between 2009 and 2016. It contributes to house price variation research in four main ways: Firstly, using the basic multilevel modelling (variance components model) to explore the house price variance in

England in two spatial scales and three different time periods. The results show that the LA spatial effect on house price is quite large, contributing 59% of total house price variance. The MSOA spatial effect within the same LA on house price variation, is relatively small with only accounting for 12%. Compared the house price difference among MSOAs, house price change within the same MSOA for different time periods (quarter, half-year, year) between 2009 and 2016 is small enough to ignore. This time influence on house price variance is the same no matter which time period in used (quarter, half-year, year). Secondly, HPM yearly trajectories in England were explored using more complex multilevel models (GCMs) due to using the year as the time period which fits the model best. Results show that the LA has a high house price in 2009 which grows relative faster over the eight-year periods than cheap LAs. Thirdly, similar to the house price ripple effect observed at regional level (Meen, 1999), HPMs at LA level largely follow this pattern. LAs in and around Bristol also show a small 'ripple effect' pattern, which maybe potentially be driven by both the London effect and by surface transport, plus a Cardiff effect via the Severn Bridge. Fourthly, spatial patterns of house price in 2009 are different to the spatial pattern of average house price percentage changes at LA in England. Using hybrid hierarchical *k*-means clustering method by considering these two difference across LAs in England, the spatial clustering pattern is observed. House price in England at LA level further reveals the ripple-effect pattern mainly driven by London, starting with extremely expensive house price high increase inner London and slowing down far away from London. This London effect shows a stronger influence in the western parts of the London rather than the east.

With a clear understanding of LA house price spatial-temporal pattern, the intention is to extend is this work through a more thorough exploration of spatial-temporal pattern of house price varation by property type. Some ONS housing affordability metrics directly estimate the housing affordability by property type without analysing its variation. Thus, the next chapter begins with an exploration of how this factor (i.e. property type) influences house price variation at LA and MSOA level between 2009 and 2016. Understanding the underlying mechanisms of house price variation in England at and below LA will not only offer deeper insights into pressing housing inequality issues, but also offer a critical suggestion on creating an effective housing affordability metric to reflect local housing affordability issues.

Chapter 6 A new insight into local housing affordability in England through further exploration of house price variation

6.1 Introduction

Buying a home is often the biggest purchase for a household over their lifetime. A large percentage of buyers in England purchase their properties with the assistance of a mortgage rather than paying directly in cash. In England between 2012 and 2019, the majority of buyers are mortgagees (Figure 6.1). During this period, 69% of buyers in England were mortgagees. For each year, over 90% of LAs in England have over 50% of mortgage buyers. Meanwhile, the number of LAs with higher than the overall England level (69%) increases significantly after 2015, from 158 in 2016 to 223 in 2019. It obvious that an increasing proportion of household in England are mortgagees.

Shocked by the GFC of 2007, house prices in England were pushed into a two-year recession between 2007 and 2009, with different rates of recovery afterwards. Chapters 4 and 5 found that England's HPM between 2009 and 2016 are found to be clustered at LA level and highly clustered at MSOA level along with a tiny temporal variation. This huge spatial house price variation in England contributes to the complexity of housing affordability issues of owning and buying a home. Chapter 5 revealed that LA's HPM between 2009 and 2016 displayed a 'fan out' trend. In such a fast house price appreciation period, affordability problems become more acute for households struggling to get on the housing ladder. However, housing affordability issues result from the interplay of price, available capital, available housing types and locational characteristics (Collinson, 2014; Galster and Lee, 2021; John, 2015). For the same property, households with different structures and characteristics will suffer different levels of housing affordability issues, which will also vary by location. These will vary spatially and socially giving rise to a complexity that can be difficult to quantify.

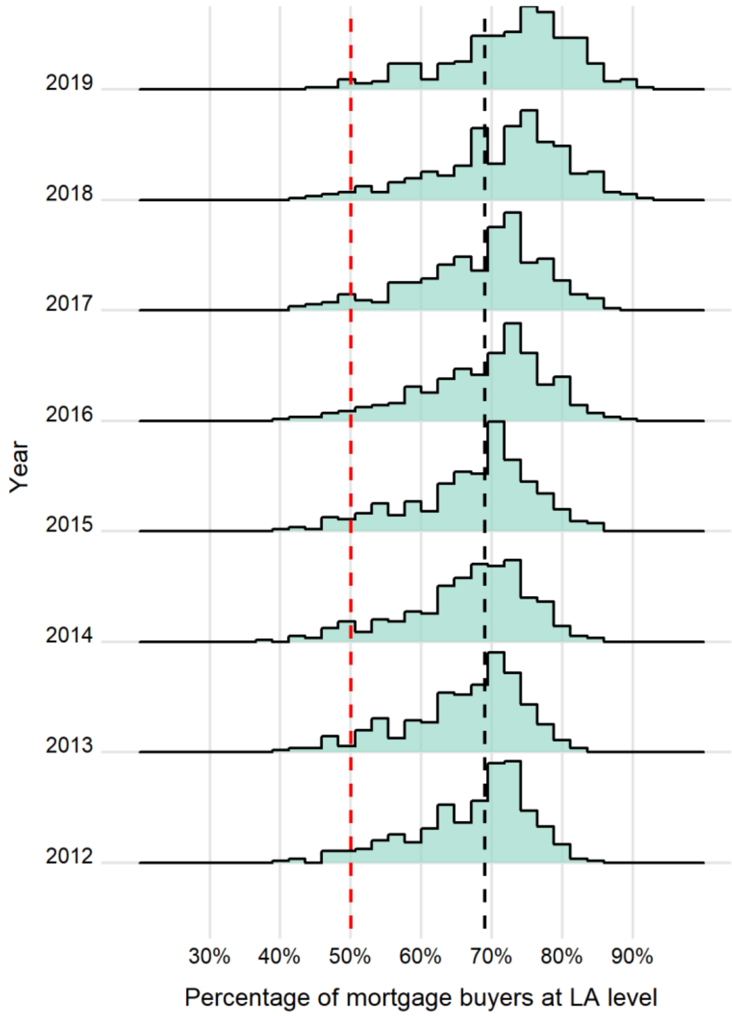


Figure 6.1 Histogram of England LA's mortgage buyers' proportion, 2012-2019[21]

The complexities within housing affordability is often crudely been measured as ratio of house price to earnings in the England (Day, 2016; Jones et al., 2011; ONS, 2017e, 2020a). Such indicators aim to illustrate the overall extent of affordability problems as

---

21 Statistic based on Cash mortgage sales in UK House Price Index (https://www.gov.uk/government/statistical-data-sets/uk-house-price-index-data-downloads-august-2020).

well as their distribution socially and geographically (Stone, 2006b). A more nuanced understanding of residential house price in England will support an in-depth understanding of the housing affordability issue. Considering housing regulation and delivery in England has been carried out by LAs since the late 19th century (Morphet and Clifford, 2020). Understanding of residential house affordability at LA level will not only help in understanding the housing affordability issue but should also assist LAs in housing-policy delivery.

The work in this chapter aims to take a similarly detailed and geographically disaggregated approach to the question of housing affordability, emulating and expanding on the ideas of the BBC calculator. Normally, households seeking to purchase a property will consider factor such as the location, property type and their available funds at the time. In this research, a novel framework is created to best reflect this reality of housing affordability, accounting for these various considerations for a particular reference household with a specific budget. In simple terms, within any given LA across England, what size and type of property can a buyer afford to buy and how has this changed over time? This novel framework is approached from two directions to overcome deficiencies in the data. The first direction focuses on maximum delineated house price variation based on our linked HPM dataset in Chapter 3. Here we consider HPM by property type at LA level to better reflect the supply side diversity with an analysis of its variation. The second direction focuses on simplifying the huge variety of possible household compositions. The huge variety of possible household compositions makes the task of assessing affordability extremely complex. What is affordable for a single young adult living alone may not be affordable for a family with several dependent children. However, neither detailed information on household composition nor on available budget for housing purchase is available in the UK. Therefore, to make the problem tractable, we create a series of typical households and study the question of affordability for each of them separately. We divide these typical households into three overlapping categories (cash buyers in England, mortgage buyers in England, and London homeowners) to support the exploration of housing affordability underlying three scenarios. The first two scenarios separately focus on offering an insight into spatial housing affordability for cash buyers and mortgage buyers underlying some designated typical households. To ease England's housing affordability driven by the London effect in terms of high HPM with relatively high price increases, the third scenario explores how housing affordability could be eased by commuting outside London. What is more, the change of housing affordability by property type for the typical-household in third scenario is further explored. The structure of this chapter is as follows. Section 2 focuses on exploration of house price variation by property type at LA level in England. It starts by introducing the data then presents the GCMs for this section. Section 2 ends with the model results along with a section summary. Section 3 focuses on the housing affordability analysis by creating an effective metric. It first introduces the datasets used in this section. It presents the creation of three typical household scenarios based on the research data and the new housing affordability metric. Housing affordability analyses using this new metric are introduced at the end of this section. Finally, we summarise and draw conclusions in section 4, alongside recommendations for future research.

6.2 HPM variation by different property types

6.2.1 HPM dataset

The HPM created in Chapter 3 is used in this chapter. From it eight fields are used, namely, TP, total floor area, HPM, transaction year, property type, MSOA codes and LA district codes. This data offers information on transaction's HPM and total floor area between 2009 and 2016. Figure C1 in Appendix C1 shows the distributions of HPM and total floor area. Both of them exhibit a positive skew with a long tail in the high value range. The majority of HPMs are below 15,000 £/m^2. The mode HPM is 2,500 £/m^2. Nearly all sold properties have a total floor area of below 400 m^2. The mode sold property size (i.e. total floor area) between 2009 and 2016 is 84 m^2.

Understanding the extent of the property size sold by different property types will benefit the following housing affordability analysis. Looking at the property size in terms of total floor area, 99.99 % of the property's size are below 400 m^2. Figure 6.2 shows the total floor area distribution below 400 m^2 by different property types in England over the research period (2009-2016). The mode value is labelled in the histogram of the four property types. It shows clearly that the most commonly sold detached residential house between 2009 and 2016 has a floorspace of 100 m^2, while this decreases to 84 m^2 for semi-detached residential houses, 78 m^2 for terrace residential houses and 60 m^2 for flats/maisonettes residential housing. Although some property types (i.e. detached and terraced) show a slightly different mode property size by each year between 2009 and 2016 (Table C1.1 in Appendix C1), the rank order among these four property types in terms of this most frequently sold size remains the same for each year. Therefore, we continue using the whole time period dataset to further explore the property size for the four property types. Minimum and maximum values of the total floor are unusual, because 557 transactions exhibit a total floor area smaller than 9 m^2 or larger than 974 m^2 and were manually removed in the previous data cleaning process. The inter-quartile range of the total floor area in detached house transactions range from 89 m^2 to 148 m^2. This indicates greater diversity in detached properties' total floor areas. Semi-detached and terrace show a similar inter-quartile. The former ranges from 72 m^2 to 102 m^2, the latter from 68 m^2 to 102 m^2. These reflect the most common property size (i.e. total floor size) for semi-detached and terraced are quite similar. Flats/maisonettes had the smallest inter-quartile spread, from 48 m^2 to 72 m^2. This reflects the modal size for sold flats/maisonettes clustering around 60 m^2. However, Figure 6.2 also shows that it is difficult to identify property type simply by the size of total floor area. For example, a property with the most common total floor area in England (i.e. 84 m^2) could be any of the four property types. Within the housing market, flats/maisonettes have the high probability of being small property while detached houses have a high probability of being large property. This contributes to the difficulty of understanding household's housing affordability by different property types.

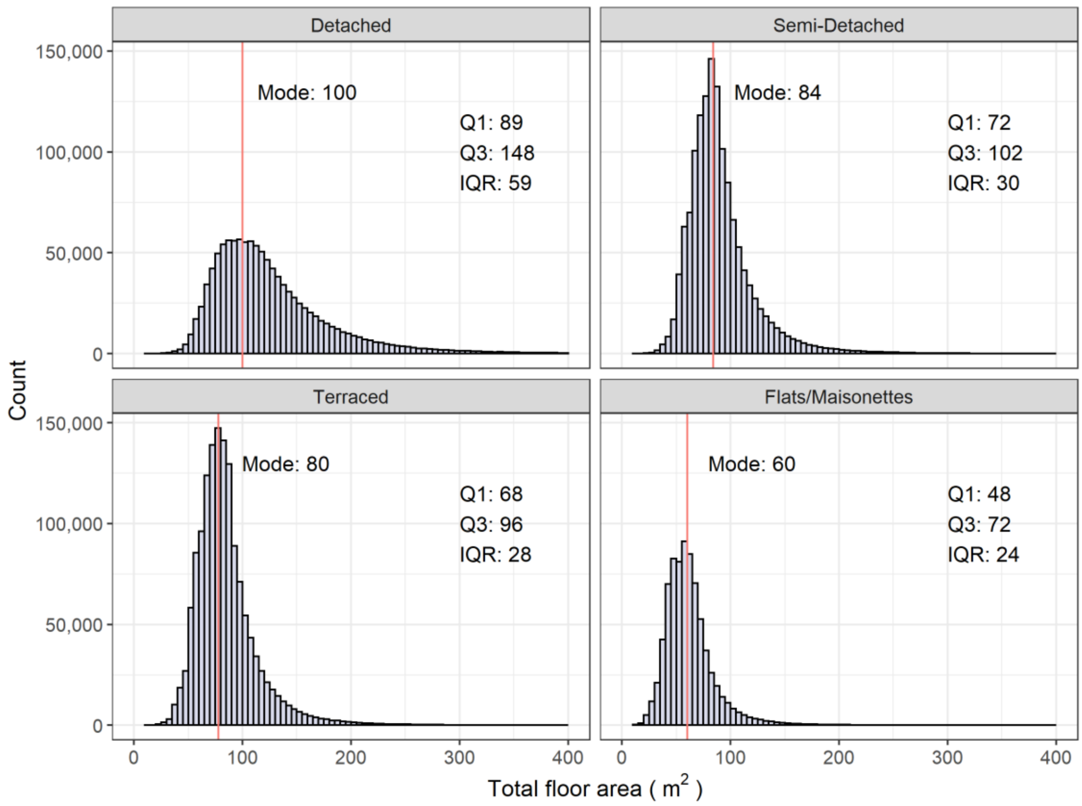


Figure 6.2 The distribution of transaction property's total floor area in England by property type[22]

Furthermore, the HPM data allows exploration of the HPM for the four property types across the LAs in England. For the 326 LAs in England, only the City of London has no semi-detached transactions between 2009 and 2016 in the original LR PPD. Additionally, there is only one detached transaction in the City of London but it failed to link with total floor area during the creation of the HPM dataset. Accepting that semi-detached and detached transaction information for the City of London is unavailable, the other LAs have all the four property types' transaction information. Table 6.1 summarises the annual sample size of HMP dataset by LA between 2009 and 2016. The majority of LAs in England have a sample size of over 30 for these four property types. For the detached, semi-detached and terraced properties, most of the

[22] Given the transactions with a total floor area over 400 m2 is relatively small (6454), this plot did not include them.

LAs with less than 30 annual sample size are located within London with one exception, i.e. Isles of Scilly in South West England. LA’s with a flats/maisonettes sample size lower than 30, are mainly in located in the East Midlands and the North West. Flats/maisonettes in the Isles of Scilly shows a similar sample size to the other three property types. LAs with lower than 30 annual sample size are listed in Table C1.2 in Appendix C1.

Table 6.1 Description summary of the annual sample size of HPM dataset by property types for LAs

Property type	**Annual sample size for LAs**				
	Min	Max	Median	Mean	Proportion LA with over 30 sample size
Detached	1	4,006	479	543.1	96.92%
Semi-detached	2	4,387	515	615.2	98.88%
Terraced	1	4,722	516	649.9	99.42%
Flats/maisonettes	1	4,684	254	445.1	96.59%

6.2.2 GCMs

MLM is a powerful method for estimating mean values using shrinkage when the sample size for a given group is small (Raudenbush and Bryk, 2002). Similar as the GCMs in Chapter 5, four sets of three-level GCMs (Table 6.2) are built to estimate the HPM variance in England for different types of property (detached, semi-detached, terraced and flats/maisonettes) at LA level between 2009 and 2016. Models D, S, T, F separately refer to GCMs for detached, semi-detached, terraced and flats/maisonettes HPM. Within the same property type model (e.g. Model D), two different types of GCM are created: Model 1 is the random intercept format, assuming all LAs have the same growth trend. Model 2 is random slope model, assuming all LAs have different growth trends. A likelihood ratio test is used to identify which format of three-level

GCM fits the data better. For each GCM, hp_{ikj} is the log HPM for a certain property type transaction i in MSOA k belonging to LA j. t_{ikj} is the year of the corresponding transaction. β_0 is overall mean house price across all LAs between 2009 and 2016. β_1 is the slope, which reflects the overall house price trend. l_j or l_{0j} is the residual at level 3, m_{kj} is the residual at level 2, e_{ikj} is the residual at level 1. l_{1j} is the random slope at level 3. The time variables (t_{ikj}) are centred at the beginning of year 2009 so that the estimated intercept has a meaningful interpretation (Raudenbush and Bryk, 2002), which refers to the estimated house price (log scale) in 2009. We refer to the estimated slope for each LA in Models D, S, T and F as "estimated house price percentage change" (e.g. $\beta_1 + l_{1j}$ in Model D2). The estimated intercept at LA level is when converted to its natural scale (i.e. exponential $\beta_0 + l_{0j}$ in Model D2) referred to as the "starting-price[23]" for each LA.

Table 6.2 The candidate three-level GCMs

Model		Equation
Model D	Model D1	$hp_{ikj} = \beta_0 + \beta_1 t_{ikj} + l_j + m_{kj} + e_{ikj}$ $l_j \sim N(0, \sigma_l^2)$ $m_{kj} \sim N(0, \sigma_m^2)$ $e_{ikj} \sim N(0, \sigma_e^2)$
	Model D2	$hp_{ikj} = \beta_0 + \beta_1 t_{ikj} + l_{0j} + m_{kj} + l_{1j} t_{ikj} + e_{ikj}$ $l_{0j} \sim N(0, \sigma_{l0}^2)$ $l_{1j} \sim N(0, \sigma_{l1}^2)$ $m_{kj} \sim N(0, \sigma_m^2)$ $e_{ikj} \sim N(0, \sigma_e^2)$
Model S	Model S1	$hp_{ikj} = \beta_0 + \beta_1 t_{ikj} + l_j + m_{kj} + e_{ikj}$ $l_j \sim N(0, \sigma_l^2)$ $m_{kj} \sim N(0, \sigma_m^2)$ $e_{ikj} \sim N(0, \sigma_e^2)$
	Model S2	$hp_{ikj} = \beta_0 + \beta_1 t_{ikj} + l_{0j} + m_{kj} + l_{1j} t_{ikj} + e_{ikj}$ $l_{0j} \sim N(0, \sigma_{l0}^2)$ $l_{1j} \sim N(0, \sigma_{l1}^2)$ $m_{kj} \sim N(0, \sigma_m^2)$

[23] All the starting-prices in this research refer to the estimated HPM in 2009 at LA level.

Model		Equation
		$e_{ikj} \sim N(0, \sigma_e^2)$
Model T	Model T1	$hp_{ikj} = \beta_0 + \beta_1 t_{ikj} + l_j + m_{kj} + e_{ikj}$ $l_j \sim N(0, \sigma_l^2)$ $m_{kj} \sim N(0, \sigma_m^2)$ $e_{ikj} \sim N(0, \sigma_e^2)$
	Model T2	$hp_{ikj} = \beta_0 + \beta_1 t_{ikj} + l_{0j} + m_{kj} + l_{1j} t_{ikj} + e_{ikj}$ $l_{0j} \sim N(0, \sigma_{l0}^2)$ $l_{1j} \sim N(0, \sigma_{l1}^2)$ $m_{kj} \sim N(0, \sigma_m^2)$ $e_{ikj} \sim N(0, \sigma_e^2)$
Model F	Model F1	$hp_{ikj} = \beta_0 + \beta_1 t_{ikj} + l_j + m_{kj} + e_{ikj}$ $l_j \sim N(0, \sigma_l^2)$ $m_{kj} \sim N(0, \sigma_m^2)$ $e_{ikj} \sim N(0, \sigma_e^2)$
	Model F2	$hp_{ikj} = \beta_0 + \beta_1 t_{ikj} + l_{0j} + m_{kj} + l_{1j} t_{ikj} + e_{ikj}$ $l_{0j} \sim N(0, \sigma_{l0}^2)$ $l_{1j} \sim N(0, \sigma_{l1}^2)$ $m_{kj} \sim N(0, \sigma_m^2)$ $e_{ikj} \sim N(0, \sigma_e^2)$

6.2.3 LA house prices change for different property types

Table 6.3 lists the results of three-level GCMs shown in Table 6.2. Each property type model shows a large deviance decrease from the first model to the second model (e.g. Model D1 vs Model D2). The related pair of likelihood Ratio tests result in an effectively zero p-value. This suggests that Model 2 fits the data significantly better than Model 1. This reveals that for any property type, the LAs' HPM growth trends vary across England.

Table 6.3 Model result of three-level GCMs

Parameter	**Model D**				**Model S**			
	Model D1		**Model D2**		**Model S1**		**Model S2**	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
β_0 Intercept	7.7302	0.0212	**7.7216**	0.0186	7.5777	0.0231	**7.5679**	0.0198
β_1 (Year-2009)	0.0317	0.0001	**0.0343**	0.001	0.0373	0.0001	**0.0405**	0.0012
σ_{l0}^2 between LA variance	0.1447	0.0115	0.1111	0.0089	0.1722	0.0137	0.1255	0.01
σ_{l01} Intercept-slope covariance	-	-	0.0039	0.0004	-	-	0.0058	0.0005
σ_{l1}^2 Slope variance	-	-	0.0003	0	-	-	0.0005	0
σ_m^2 between MSOA variance	0.0219	0.0004	0.0221	0.0004	0.0282	0.0005	0.0283	0.0005
σ_e^2 Individual variance	0.0607	0.0001	0.0597	0.0001	0.0626	0.0001	0.0605	0.0001
Deviance	64,146		47,531		121,653		76,664	
Parameter	**Model T**				**Model F**			
	Model T1		**Model T2**		**Model F1**		**Model F2**	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
β_0 Intercept	7.4829	0.027	**7.4829**	0.0226	7.4162	0.024	**7.4998**	0.0186
β_1 (Year-2009)	0.0401	0.0001	**0.0403**	0.0015	0.051	0.0001	**0.0294**	0.0017
σ_{l0}^2 between LA variance	0.2341	0.0185	0.1634	0.013	0.1851	0.0147	0.1097	0.0088
σ_{l01} Intercept-slope covariance	-	-	0.008	0.0008	-	-	0.0078	0.0007
σ_{l1}^2 Slope variance	-	-	0.0007	0.0001	-	-	0.0010	0.0001
σ_m^2 between MSOA variance	0.0399	0.0007	0.0403	0.0007	0.0399	0.0008	0.0396	0.0008
σ_e^2 Individual variance	0.0723	0.0001	0.0682	0.0001	0.0816	0.0001	0.0762	0.0001
Deviance	334,602		251,629		277,561		226,826	

Comparing the overall mean HPM (β_0) in England across the four property types, detached is the most expensive property type (exp(7.7216) = £2256.57 per metre squared) followed by semi-detached (exp(7.5679) = £1,935.07 per metre squared). The overall mean HPM for flats/maisonettes (exp(7.4998) = £1,807.68 per metre squared) and terraced (exp(7.4829) = £1,777.39 per metre squared) are quite similar, but the flats/maisonettes show a marginally higher overall mean HPM than terraced. However, the overall house price change (β_1) shows a different rank order from the overall mean HPM among the four property types. Semi-detached houses showed the biggest overall house price increase (in percentage terms) between 2009 and 2016 followed by terraced houses. Detached houses, the most expensive starting-price category, saw a relatively less steep increasing trend. Flats/maisonettes showed the lowest increases over the period.

Covariance between the intercept and slope in Models D, S, T and F are positive, suggesting a positive relationship between the LA house price increase and starting-price within the same property type. It reveals that the HPM for each property type follows a similar 'fanning out' of growth trends at LA level in England for each year between 2009 and 2016. Below we separately explore this fanning out trend from two dimensions. One is from the starting-price (the intercept), the other is from the overall HPM percentage change (the slope).

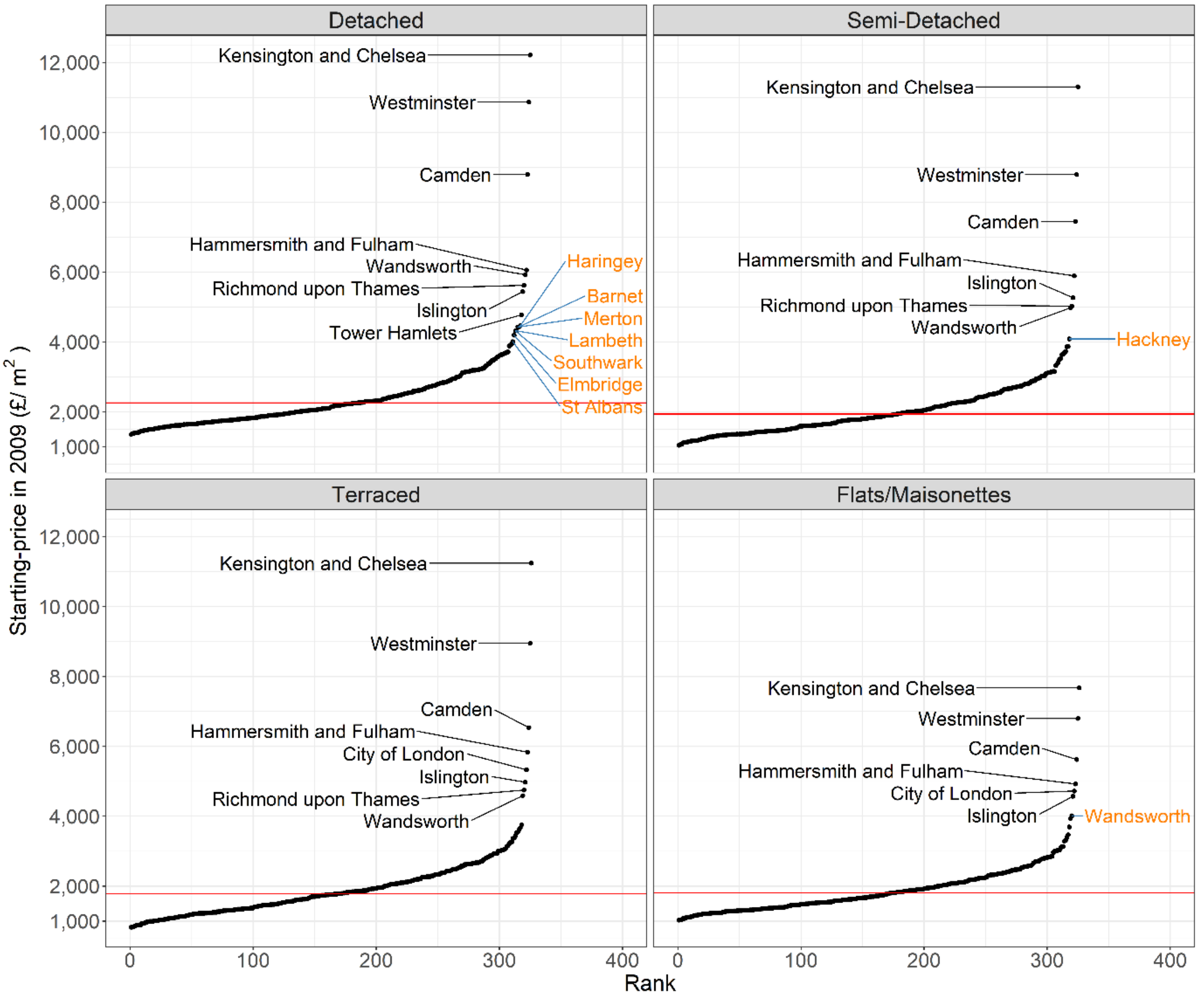


Figure 6.3 Starting-price at LA level for four property types

Figure 6.3 displays the estimated starting-price in 2009 for different property types among LAs in England. Each point represents one LA in England and the red horizontal line shows the overall starting-price in 2009 in England. Looking at the average starting-price in England for detached, demi-detached, terraced and flats/maisonettes, we see that the detached starting-price (2,257 £/m^2) is most expensive, followed by semi-detached (1,935 £/m^2) and then flats/maisonettes (1,807 £/m^2) followed by terraced (1,777 £/m^2). Figure 6.3 also labels the LAs with prices over 4,000 £/m^2. It obvious that these labelled LAs are the key contributors to the big HPM variation at LA level in England. The remaining LAs all have HPMs below 4000 £/m^2.

The top three most expensive LAs in England for different property types are the same. They are Kensington and Chelsea, Westminster and Camden. Excepting the flats/maisonettes HPM in Camden, all the other property types' HPM in these three LAs maintain HPM levels higher than 6,000 £/m^2. Kensington and Chelsea show the biggest HPM differential among these four property types followed by Westminster and Camden. Detached is the most expensive HPM property type in these three LAs, and is very much higher than for the other three property types (semi-detached, terraced and flats/maisonettes). This leads to the large variation of the property size among these four property types with the same property value when estimating the affordable property size. This situation is also replicated in all the other labelled LAs in Figure 6.3. Those LAs with an average HPM of over 4,000 £/m^2, exhibited a large HPM difference in property types. This shows that HPM variation within LAs is affected by the makeup of property types in the areas.

Figures 6.4 to 6.7 separately plot the spatial patterns of LA starting-prices for the four property types. These four choropleth maps use the equal 1,000 £/m^2 interval as in Figure 5.6. For each map, the LAs with HPM between 4,000 £/m^2 and 6,000 £/m^2 are labelled. The majority of LAs in England have an HPM below 3,000 £/m^2 (LAs shaded in blue). LAs with HPM over 3000 £/m^2 are mainly located in London or nearby. For the detached property type, 59 LAs have an HPM over 3,000 £/m^2. Excepting Cotswold, Cambridge, Winchester, Chichester, Horsham, and Brighton and Hove, the rest of the 59 LAs are mainly located in London or nearby. For the semi-detached property type, 31 LAs have an HPM of over 3,000 £/m^2. Apart from Oxford, the rest of these LAs are mainly located in London or nearby. For the terraced property type, 28 LAs have an HPM of over 3,000 £/m^2. Except for Cotswold, Cambridge and Oxford, the LAs are mainly located in London or nearby. For the flats/maisonettes property type, 20 LAs have an HPM of over 3,000 £/m^2. Except for Oxford, the rest of the LAs are located in London or nearby. For the LAs in London, those in the West maintain a higher HPM than those in the East.

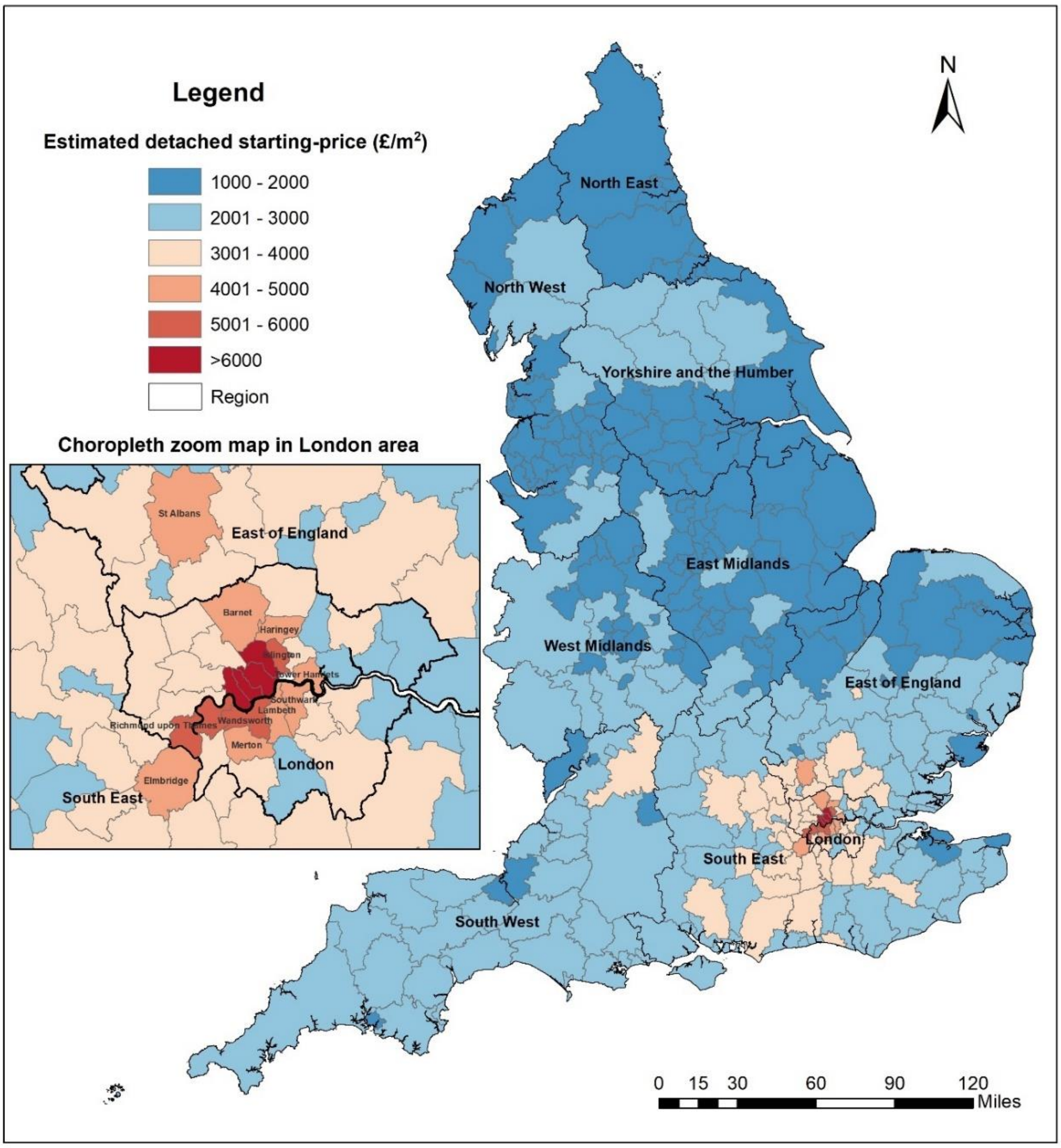


Figure 6.4 The spatial patterns of LA detached starting-price in 2009

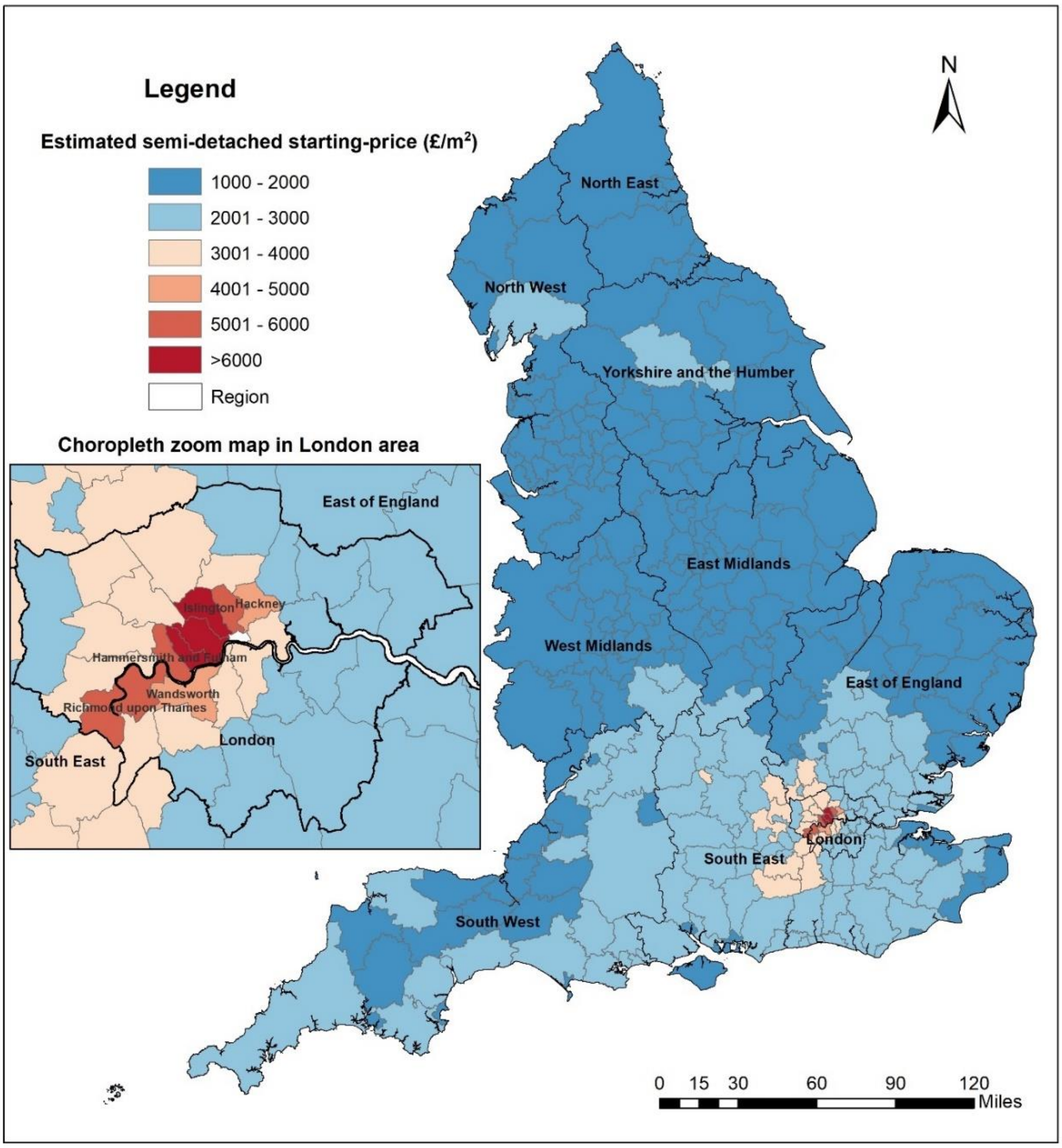


Figure 6.5 The spatial patterns of LA semi-detached starting-price in 2009

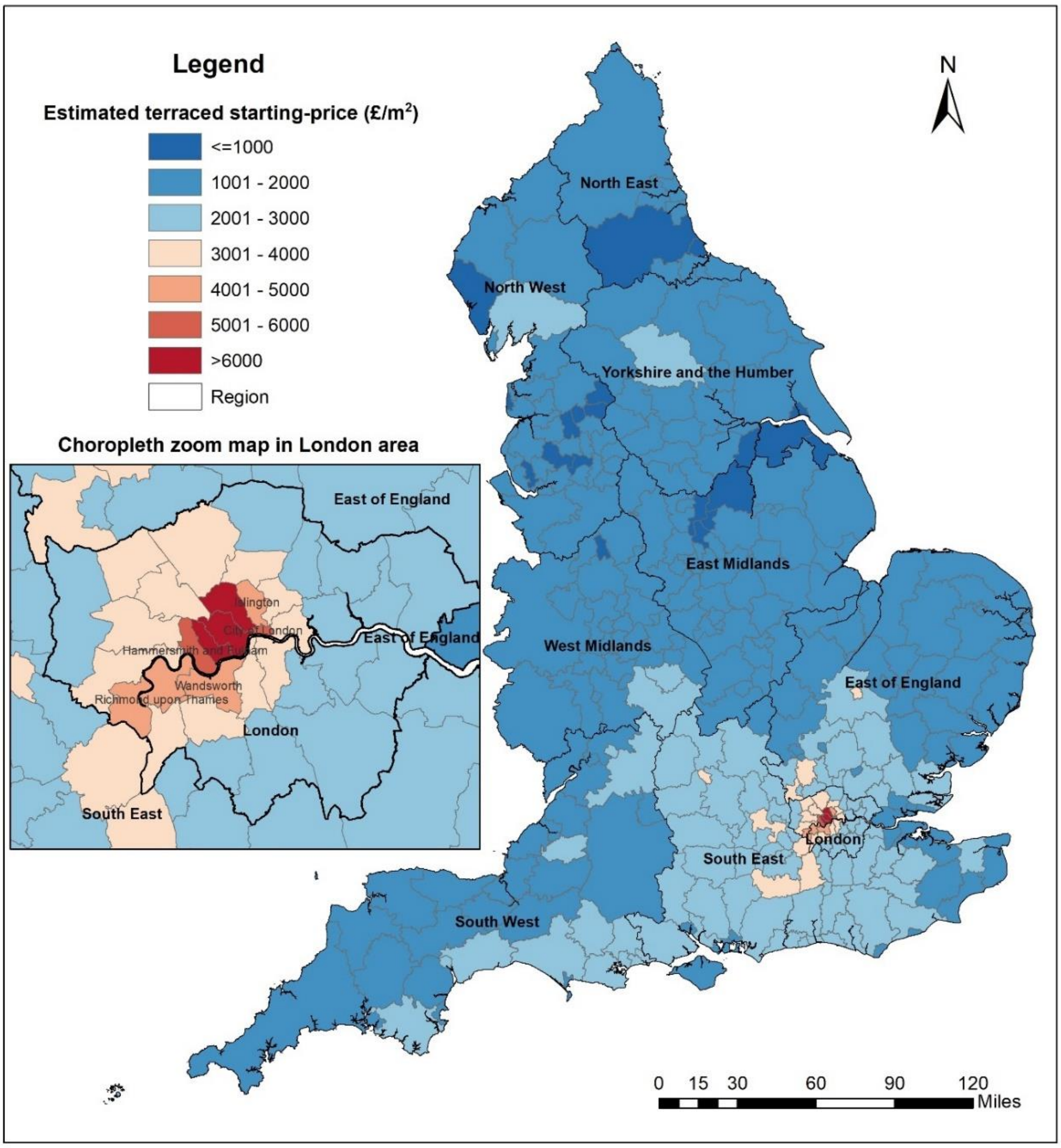


Figure 6.6 The spatial patterns of LA terraced starting-price in 2009

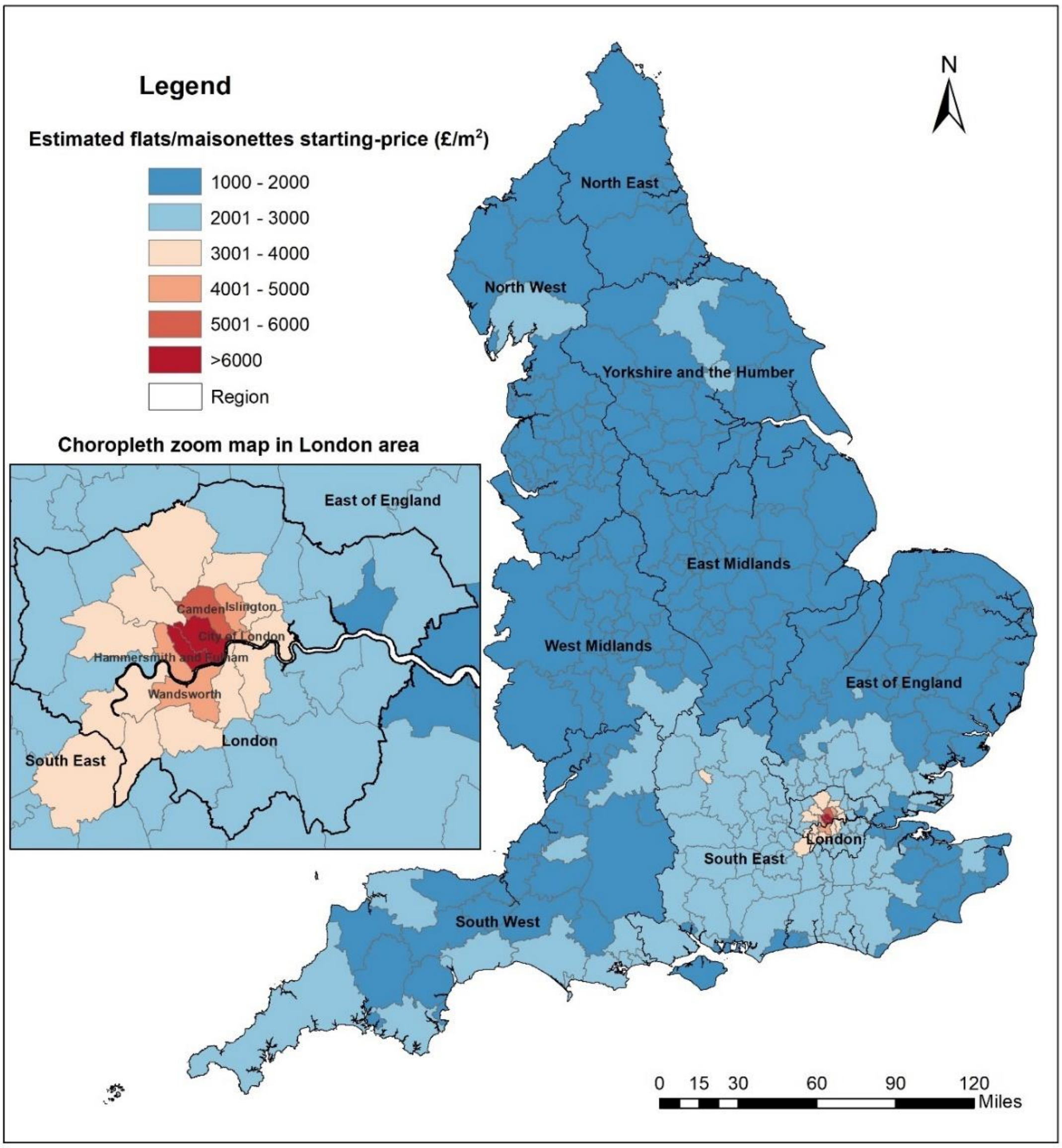


Figure 6.7 The spatial patterns of LA flats/maisonettes starting-price in 2009

Unlike the wide spread of LAs' starting-prices, the overall LAs' house price change is more concentrated (Figure 6.8). As in Figure 6.3, each point in Figure 6.8 represents one LA in England and the red horizontal line presents the overall HPM percentage change in England. Generally, the HPM in England for the four different property types displays an increasing trend. The majority of LAs show an overall HPM increase of under 10%. Figure 6.8 labels the LAs with an overall HPM increase above 10%. For a given property type, LAs with the most expensive starting-price did not have the highest percentage HPM increase. For example, those LAs with the highest starting-price for detached and semi-detached houses, did not show the biggest overall percentage HPM increase for those property types. Instead, terraced HPM and flats/maisonettes HPM in Hackney, Lambeth, Lewisham, Waltham Forest and terrace HPM in Southwark show the biggest HPM increases between 2009 and 2016, at over 10%.

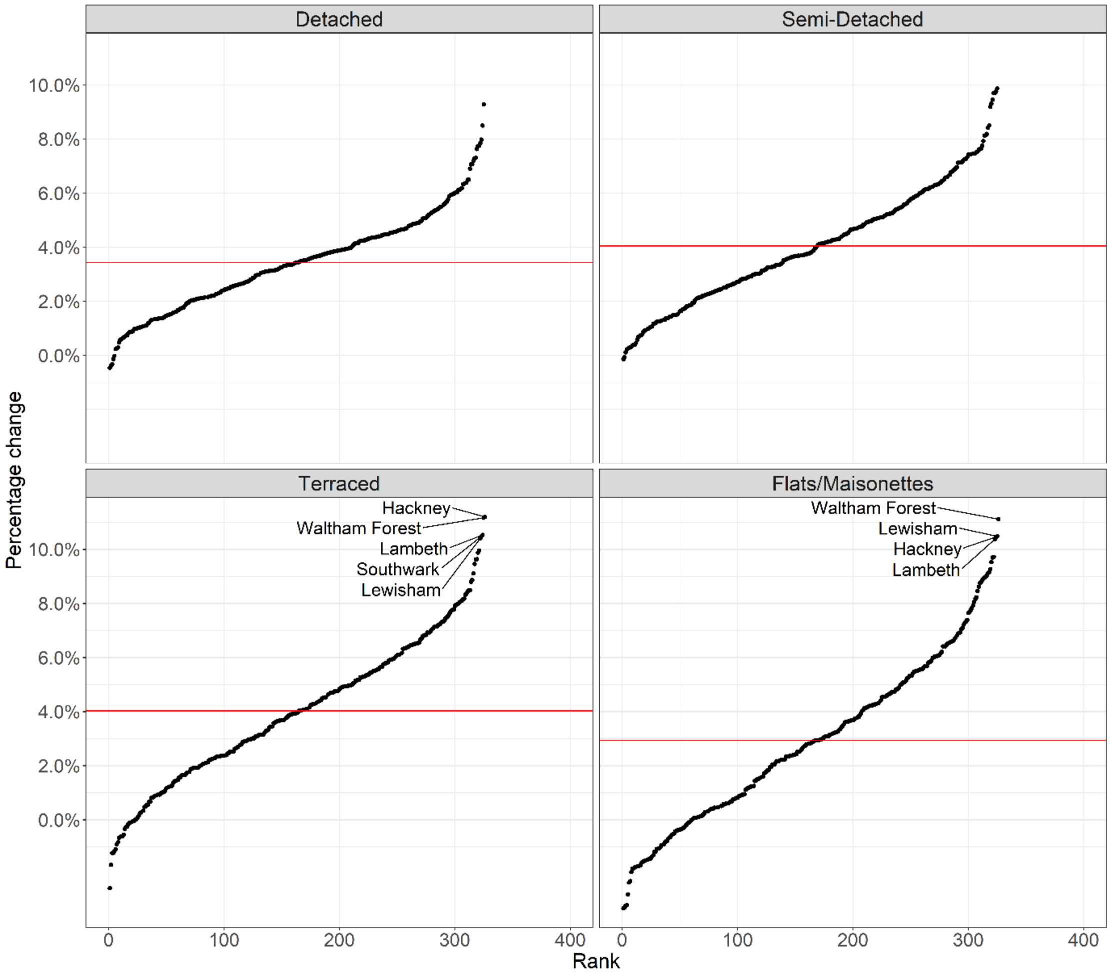


Figure 6.8 House price change at LA level for four property types

Looking at the spatial pattern the overall HPM percentage change at LA level for the four property types (Figures 6.9 to 6.12), follows a similar radial gradient pattern to that of the house price change in England (Figure 5.9 in Chapter 5). There are two obvious gradient patterns of percentage change at LA level. One is centred on London and the other is centred on Bristol. Over half of the LAs have a HPM percentage change below 4%. Flats/maisonettes HPM and terraced HPM in some LAs represent a small decreasing trend. All property types' HPM, display a high increase in London or nearby. Within London, each property type HPM at LA level generally shows a percentage price increase greater than 6%, but the detached HPMs in outer London show a percentage increase lower than 6% but higher than 4%.

The spatial pattern of HPM percentage change for semi-detached properties is quite similar to that for terraced properties. Almost 50% of LAs show HPM percentage increases over 4% whereas the proportion of LAs with over 4% increase for flats/maisonettes and detached houses drops to 35%. Similarly, comparing the HPM change pattern for detached, semi-detached and terraced house, flats/maisonettes in the North of England show the lowest level of price change.

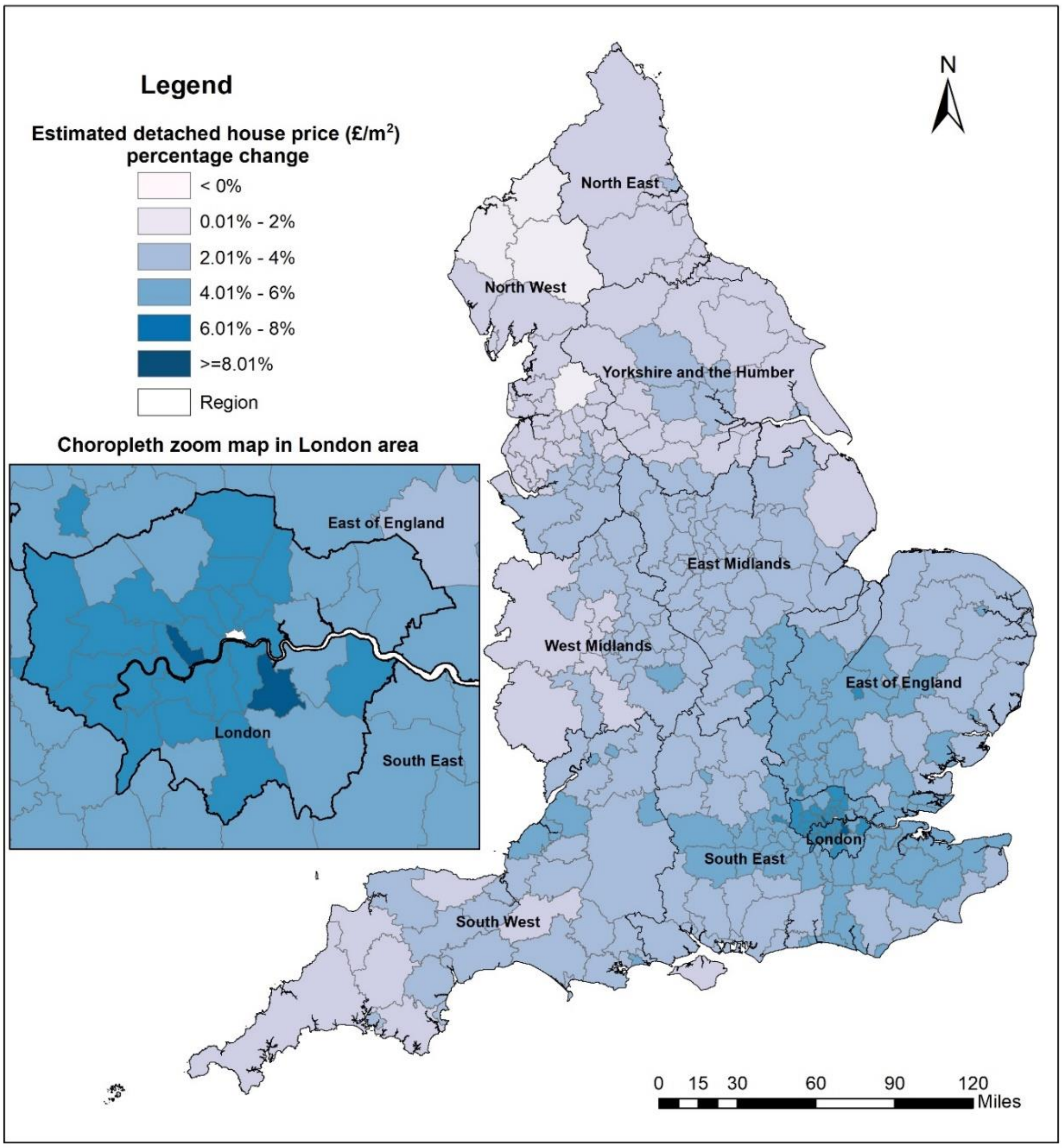


Figure 6.9 The spatial pattern of average detached HPM percentage change at LA level

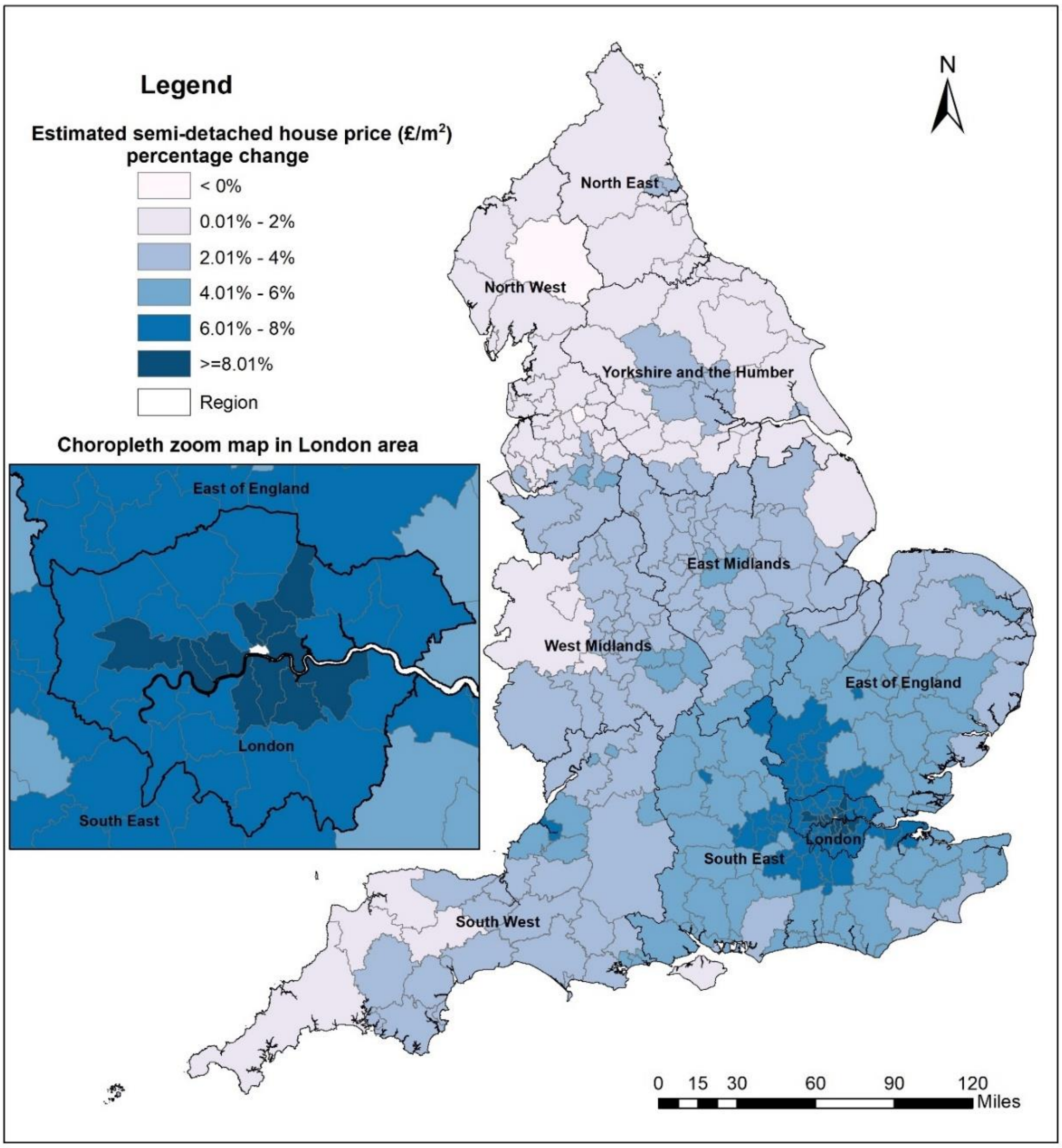


Figure 6.10 The spatial pattern of average semi-detached house prices percentage change at LA level

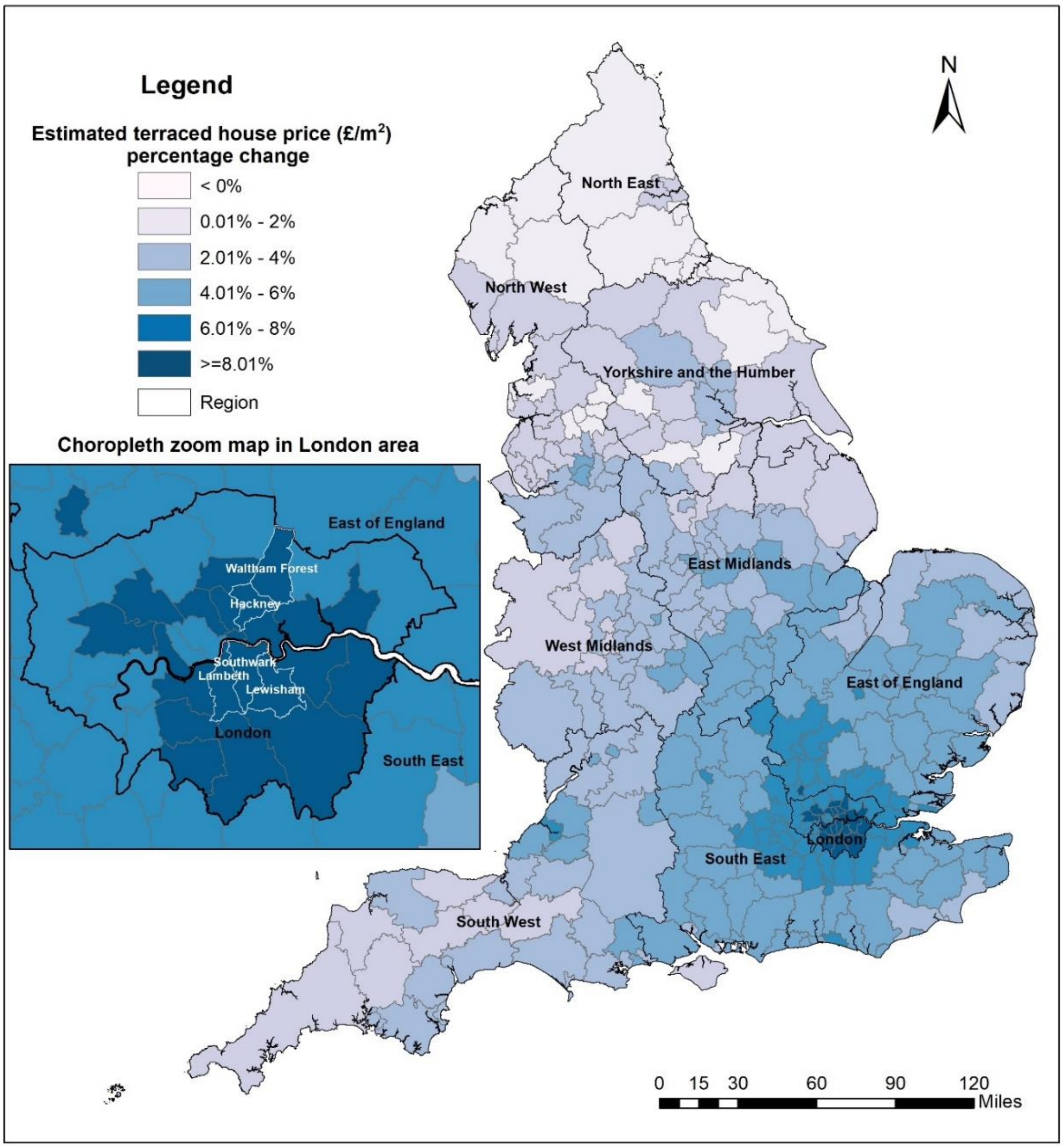


Figure 6.11 The spatial pattern of average terrace house prices percentage change at LA level

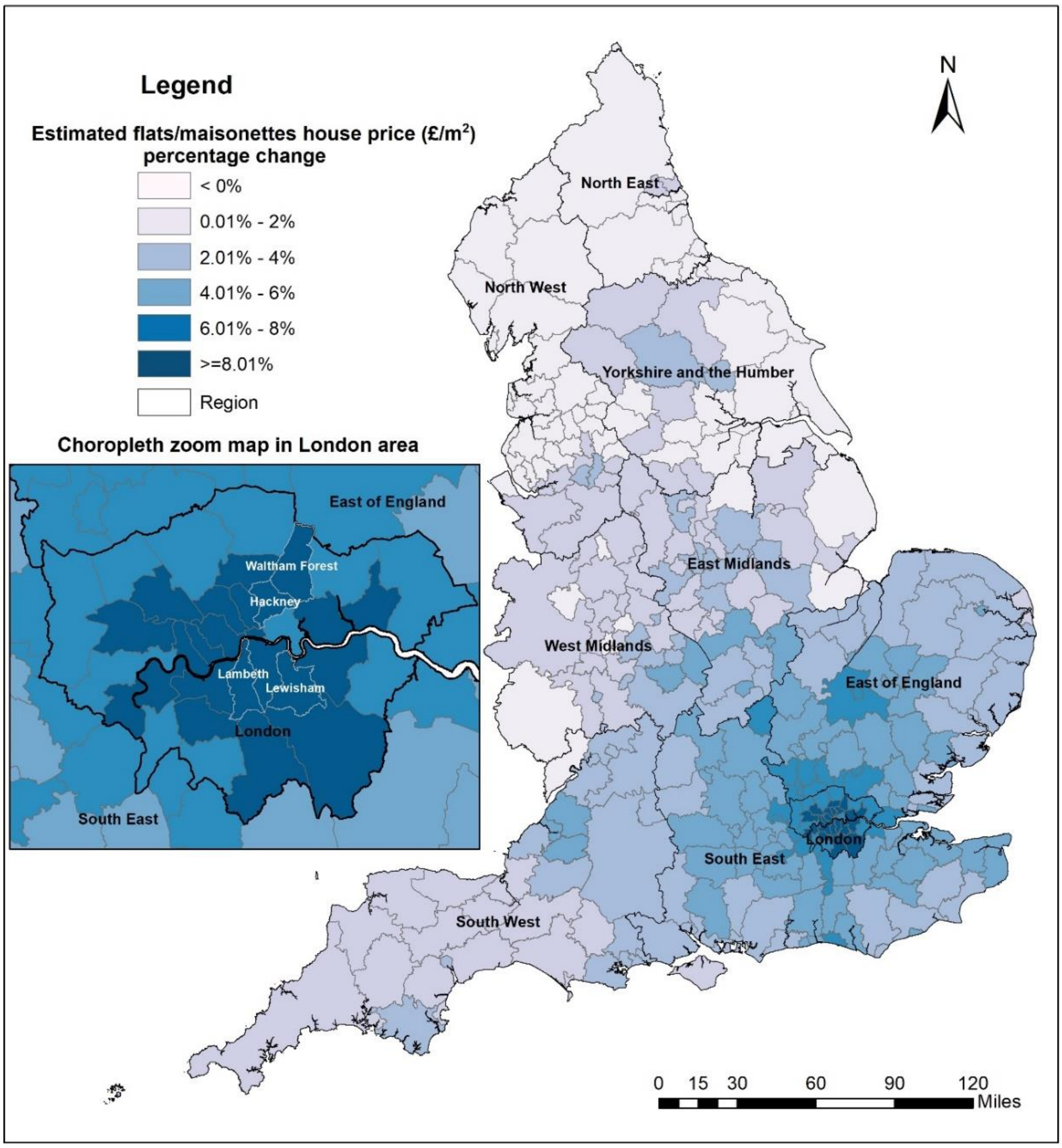


Figure 6.12 The spatial pattern of average flats/maisonettes house prices percentage change at LA level

6.2.4 Section discussion

The spatial patterns for starting-price (HPM) or overall HPM change are similar among property types, but they vary in detail. Among the four different property types, the spatial patterns of the starting-price (HPM) generally show a high HPM in London (Figure 6.4 to Figure 6.7) and a lower HPM further from London. However, there are more LAs in London with a detached HPM over 3,000 £/m^2, than with an HPM over

3,000 £/m^2 for the other three property types. LAs in London have a high level starting-price, which is similar for semi-detached and terraced housing types, but there is a different pattern outside London. The number of LAs in London with high HPM level for flats/maisonettes are the smallest. Starting-prices (HPM) at LA level within London show greater variation than LAs outside London. The majority of LAs outside London show house prices lower than 3,000 £/m^2 with the HPM for detached properties being relatively higher than for the other three property types. This not only means starting-prices are different in property types at LA level in England, but also means that housing affordability for the same household will show a larger difference among property types within London rather outside London. The starting-price difference among property types and these differences across LAs reveal the necessity of understanding housing affordability by property type.

The spatial patterns of overall HPM change by different property type (Figure 6.9 to 6.12), show some difference between different LA's neighbourhoods. Similarly, to the spatial pattern in starting-price, the spatial pattern for semi-detached and terrace house in London LAs is almost the same but differs outside London. However, LAs with the highest HPM across the different property types did not represent the highest percentage increase between 2009 and 2016. The difference in starting-price and overall house price increase by different property types are not the same. This further contributes to the complexity of assessing housing affordability by property types across the space and time.

6.3 Housing Affordability analysis

6.3.1 Data

6.3.1.1 TP data

Considering that the HPM dataset represents 80% of the full housing market value sales, it could cause potential bias when used in reflecting the real total TP within England. Therefore, the TP from the original LR PPD between 2009 and 2016 underlying full market sales is used. It is created by using the original LR PPD in England and Wales between 2009 and 2016 to remove three types of transactions. The transactions removed all those in Wales, then those residential properties not sold at full market value or those whose property postcode no longer exists in the NSPL2017 datasets. There are 5,865,856 cleaned transactions in England left to support the following analysis of transaction costs for households between 2009 and 2016. 595,185 of them are the transactions in 2009. The bottom 95% of TPs in 2009 are under £500,000. Figure C2.1 (Appendix C2) represents the TP distribution in 2009 of this bottom 95%. The descriptive statistics for the whole dataset are labelled in the figure. It is clear that half of the transactions are below £170,000 in value. The general TP follows a positively skewed pattern with two significant peak bars. The two peak bars are located at the two SDLT bands, namely £175,000 and £ 250,000. In 2009, SDLT at 1% was required on transactions between £175,000 and £250,000. The underlying reason for a big transaction volume occurring below £ 175,000 is that home buyers were avoiding those property transactions exceeding £175,000. 12.77% of England housing transactions are in London, which contributes 21.41% of total TP. Figure C2.2 (Appendix C2) displays the 2009 TP distribution in London for the lower

95%. It uses the same price range as in Figure B1. It also labels out four basic statics based on the whole cleaned TPs in London. The TP in London distribution shows the same two peaks at £ 175,000 and £ 250,000, but it has a higher median value than the median value in England. The median transaction value in London for 2009 is £250,000, which is equal to its modal value.

6.3.1.2 Household dataset

The household dataset used in this research comes from the EHS between 2008 and 2017. It is a continuous national survey first commissioned by the Ministry of Housing, Communities and Local Government (MHCLG) in 2008. It contains two separate databases, Household Data and Housing Stock Data. The Household Data records household information on age, income, housing cost (i.e. mortgage payments), tenure and regional location etc. based on individual interviews in England. Table 6.4 summarises the total household sample size and the sample size for different tenures in the household dataset for the periods between 2008 and 2016. In this period EHS surveyed over 100,000 households in England with a slightly decreasing sample size between 2008 and 2016. The sample size of over 17,000 is maintained before 2010 after which it declines to around 13,000. The majority of households surveyed are home owners, comprising over 56% of the overall sample.

Table 6.4 A summary of the sample size of EHS (Household Data), 2008-2017

Household Data	**Total**	**Tenure**				
		Home ownership			**Private rented**	**Social rented**
		Own outright	Buying with mortgage	Total (Proportion)		
EHS, 2008-2009: Household Data	17,691	5,954	6,314	12,268 (69.35%)	2,223	3,200
EHS, 2009-2010: Household Data	17,042	5,672	5,950	11,622 (68.20%)	2,331	3,089
EHS, 2010-2011: Household Data	17,556	6,107	5,930	12,037 (68.56%)	2,470	3,049
EHS, 2011-2012: Household Data	13,829	4,271	4,288	8,559 (61.89%)	2,079	3,191
EHS, 2012-2013: Household Data	13,652	4,161	4,119	8,280 (60.65%)	2,103	3,269
EHS, 2013-2014: Household Data	13,276	3,996	3,773	7,769 (58.52%)	2,058	3,449
EHS, 2014-2015: Household Data	13,174	4,134	3,683	7,817 (59.34%)	2,087	3,270
EHS, 2015-2016: Household Data	13,468	4,205	3,543	7,748 (57.53%)	2,061	3,659
EHS, 2016-2017: Household Data	12,970	3,995	3,312	7,307 (56.34%)	2,507	3,156

The Household Data in EHS are stored in two separated datasets, generalfsxx and interviewfsxx. Together they record over 120 variables relating to household circumstances. Appendix C3 lists 39 core variables which are used in this research. Some variables change their name between 2008 and 2016. Starting from 2014, some detailed variables are only available through the Special Licence (SL). For example, *Buypresh*[24] (Year household reference person bought present accommodation) is available under SL after 2014. This SL constraint limits the ability to identify the buyers in a given year (e.g.2009). Identifying buyers in a given year will result in a more accurate measurement of housing affordability. Influenced by the increasing house price after GFC, housing affordability for a certain types of home buyers (e.g First-time buyer) could differ by years. Thus, we only use the EHS before 2014 in this research.

Using *ftbuyer, tenure2*, and *Buypresh* variables (Table C3 in Appendix C3), we are also able to differentiate the home ownership household by tenure, purchase year for first-time buyer or not. As the HPM dataset in this research starts from 2009, we only consider households who bought a house in the same year. Table 6.5 lists a summary of sample sizes of 2009 home buyers in two categories (own outright, buying with mortgage-including shared ownership) for first-time buyer or not. Overall, there are 455 own outright buyers and 1,038 mortgage buyers in 2009. Within these buyers in 2009, 76 own outright buyers are first-time buyers as are the 433 mortgage buyers. Given the household linkages between different annual EHS are unavailable, we assume that the 2009 home buyers surveyed in different annual EHS are not the same.

[24] Text written as italic refers to variable name

Additionally, we also assume that the sample size for owning outright, buying with a mortgage, first-time buyer owning outright, first-time buyer through a mortgage, non-first-time buyer owning outright, non-first-time buyer through a mortgage are big enough to represent the related buyers in 2009.

Table 6.5 A summary of sample size of 2009 home buyers in EHS, 2008-2014

Household Data	**Own outright (2009)**		**Buying with mortgage[25] (2009)**	
	Total	First-time buyer	Total	First-time buyer
EHS, 2008-2009: Household Data	1	1	2	1
EHS, 2009-2010: Household Data	59	12	155	59
EHS, 2010-2011: Household Data	136	20	274	113
EHS, 2011-2012: Household Data	89	15	233	103
EHS, 2012-2013: Household Data	82	15	207	93
EHS, 2013-2014: Household Data	88	13	167	64
Total	455	76	1,038	433

Owning outright buyers own the house directly through paying for the property upfront with cash. Mortgage buyers will fully own their house when they pay off the mortgage. The *Mortwkx* variable shows weekly mortgage payments of mortgagees. Figure 6.13 represents the weekly mortgage payments distribution for first-time buyers and non-first-time buyers. Since the weekly mortgage payments are positively skewed, the median value of weekly mortgage payments is used to reflect the average weekly mortgage payments for first-time and non-first-time buyers. First-time buyers' weekly mortgage payments are generally smaller than those for non-first-time buyers. The majority of 2009 first-time buyer mortgagees paid £130.38 per week, while the non-first-time buyers paid £161.54 per week. Comparing the range of weekly mortgage

[25] Buying with mortgage also including shared ownership

payments of these two groups, the first-time buyer payment showed a relatively small range with low repayment levels. Almost 38% of the first-time buyer's weekly mortgage repayments are between £100 and £180.

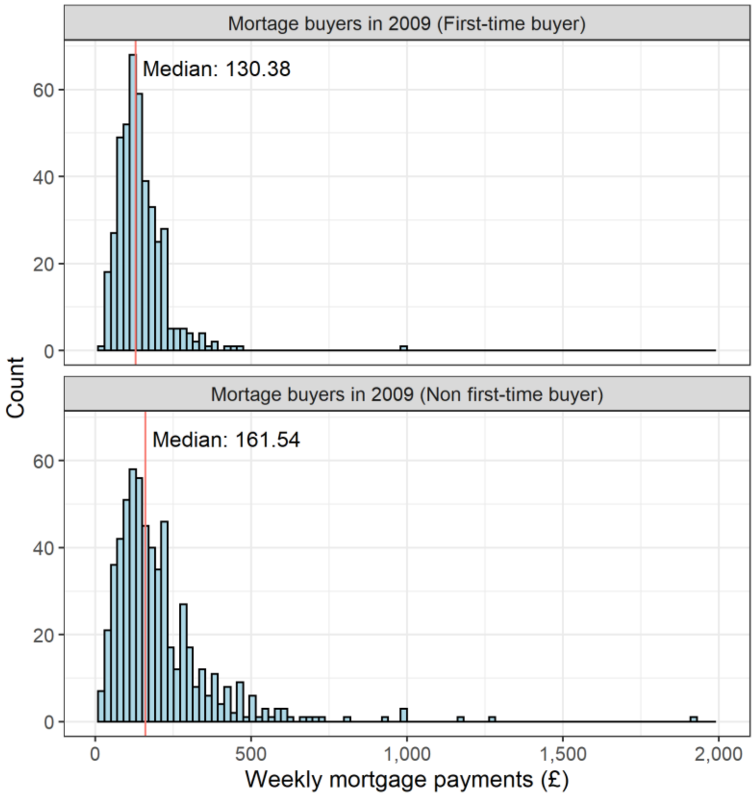


Figure 6.13 Weekly mortgage payments distribution for mortgage buyers in 2009

6.3.1.3 Commuter flow data

Census 2011 travel to work data is used to identify workers' who commute to London from outside. The Census travel to work data (Table WU03EW_MSOA) is aggregated at LA level and then we treated all the LAs in London as one unit. The proportion of extra-LA commuting that goes to London is assessed using the number of people commuting outside of home LA to work in London divided by the number of people commuting to work beyond their home LA.

6.3.1.4 Commuting time to London by public transport at LA level

Commuting time to London is assessed using the reachable areas data for each railway station in England estimated at 7 a.m. for a weekday. These are areas reachable from each station within 15 minutes, 30 minutes, 45 minutes, 60 minutes, 75 minutes, 90 minutes, 105 minutes and 120 minutes by all public transport modes and were extracted from the TravelTime platform (https://www.traveltimeplatform.com/) for 18th October 2019. Since there is no up to date integrated spatial dataset for English railway stations and routes, it was newly created by using the OS VectorMap District (2018) as a base map and then the stations and rail route information were manually checked from the GB Rail General Transit Feed Specification (GTFS) data (2016)[26], estimates of station usage 2017-18 (11 December 2018 version)[27] and the national rail route diagram map (June 2019) [28], national rail train operators map (September 2019) [29], the London Connections Map[30] and the London Tube Map (December 2019) [31]. The railway station data in OS VectorMap District covers light rapid transit stations, railway stations, and London underground stations. Given the OS stations data does not reflect up to date railway station spatial data, stations which were not present in the 2017-18 estimates of station usage are removed. Any remaining stations not shown in the GB Rail GTFS, the national rail route diagram map or the national rail train operators map were further removed manually. This left 2,267 railway stations for use

---

[26] Resource: CASA QUANT

[27] Resource: https://dataportal.orr.gov.uk/statistics/usage/estimates-of-station-usage/

[28] National rail route diagram map resources: https://www.nationalrail.co.uk/stations_destinations/maps.aspx

[29] national rail train operators map resources: https://www.nationalrail.co.uk/stations_destinations/maps.aspx

[30] Resources: https://www.whatdotheyknow.com/request/224813/response/560395/attach/3/London%20Connections%20Map.pdf

[31] Resources: the December 2019(b) version: http://content.tfl.gov.uk/large-print-tube-map.pdf

in the analysis,[32] of which 594 (26%) are located in London. The railway routes are created by merging railway tracks and tunnels from OS VMD and then manually deleting all the routes not mapped in the national rail route diagram and the national rail train operators and London tube maps. The spatial tube routes in London are derived from the London Connections Map. This newly created railway station and railway route spatial datasets are shown in Figure 6.14.

Based on the stations mapped out in Figure 6.14, the TravelTime tool (TravelTime plugin in QGIS) is used to get the reachable areas for each station at 7 a.m., within 15 minutes, 30 minutes, 45 minutes, 60 minutes, 75 minutes, 90 minutes, 105 minutes and 120 minutes by public transport. Figure C4 (Appendix C4) offers one example of the above reachable areas at Brighton train station. Areas reachable within 15 minutes are removed in the following analysis as most LAs outside London are not within 15 mins of London. Spatial joins between the rest of reachable areas and the London boundary are conducted to identify the shortest commuting time to London for each station. For the stations located in the same LA, the minimum commuting time from these stations is used to represent the LA's commuting time to London.

[32] For the station which has multiple entrances we will only keep one record.

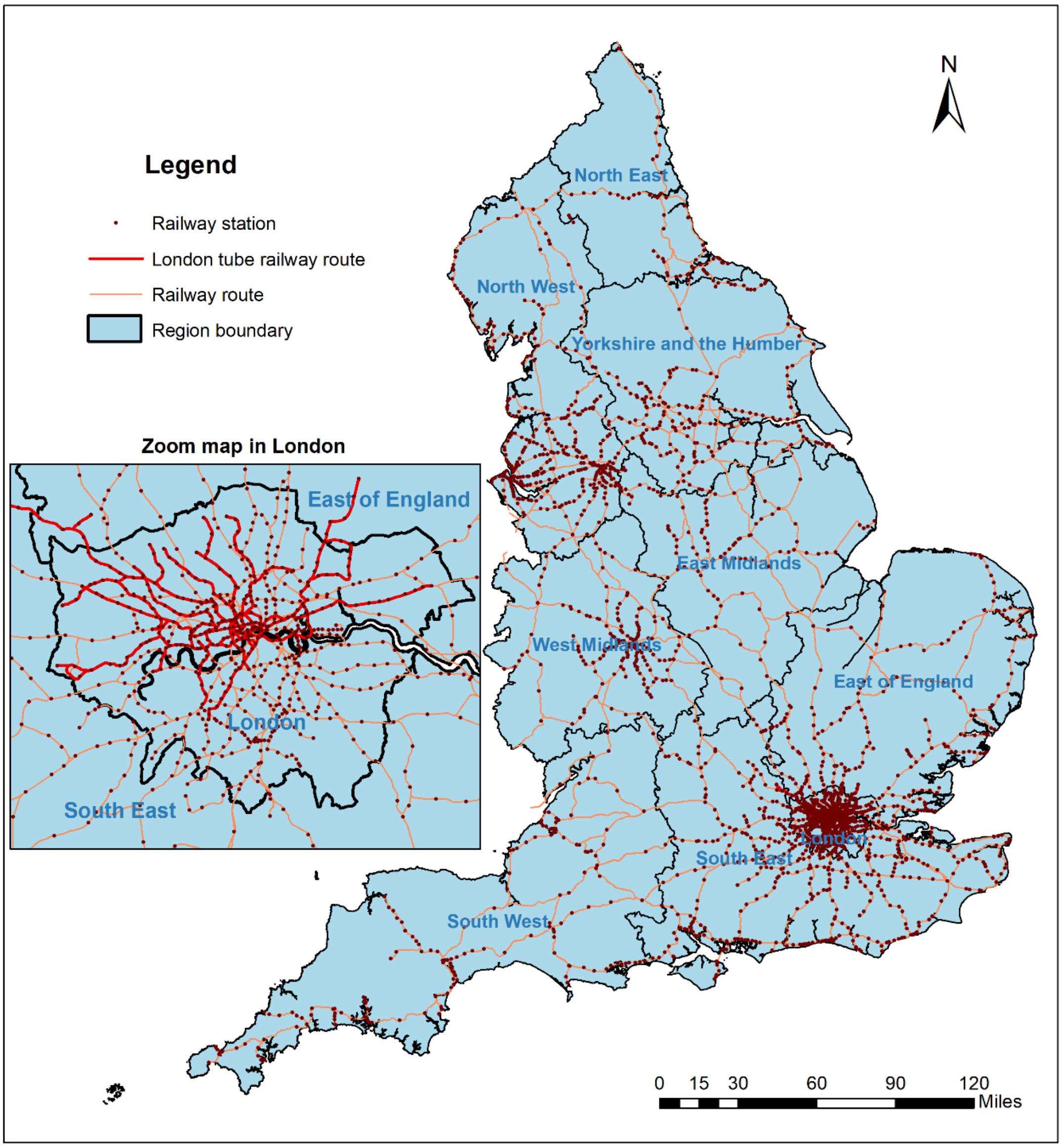


Figure 6.14 Railway stations and railway routes in England

6.3.2 Methodology

6.3.2.1 Three household scenarios in housing affordability measurement

In order to take into account the diversity of property buyers in the UK housing market, this research created three typical household buyer scenarios. The details of the household circumstances designed into the three scenarios are listed in Table 6.6. The first scenario (scenario A) is a typical household who bought housing by cash in a given year. This typical buyer is defined as buying the house with a budget equalling the median TP for the given year. The second scenario (scenario B) focusses on mortgage buyers. Two typical households are considered in this scenario, one is a first-time buyer, the other is a non-first-time buyer. They are assumed to have the median weekly mortgage repayment value for the given year according to EHS household data. The deposit size paid and mortgage interest rates can vary. For simplicity, three common deposit sizes (i.e. 5%, 10% and 25% of the property value) are selected , which is consistent with the mortgage housing affordability measure in ONS (ONS, 2020c). The monthly mortgage interest rates are chosen as the average monthly secured loan (mortgage) rates on 2 year fixed-rate mortgages for the same year as set out by the Bank of England (Bank of England, 2020). The interest rate for different loan to value (LTV) mortgages did vary, lower LTV mortgages normally have relative high mortgage interest rate. Since the monthly interest rate for 95% LTV and 90% LTV is not available or is incomplete, the average 75% LTV monthly interest rate in 2009 is used as the default interest rate in the second scenario. In this second scenario, the standard mortgage term in the UK (25 years) is used (Jones, 2017). This is also under the maximum mortgage terms (35 years) in the Bank of England's mortgage modelling. The third scenario (scenario C) models a typical home mover who already owns a property in London and wishes to move outside to buy a bigger property in a given year. London with its extremely high house prices (compared to England as a whole), may increasingly push Londoners to move out to a larger but more affordable property (Andrews, 2020; Felton, 2018). In scenario C, the typical household is assumed to possess a budget with the median property value in London in a given year and wishes to find a bigger property within the same year.

Table 6.6 A list of household circumstances designed in three scenarios for housing affordability in 2009

Scenario ID	Buyer type	Type ID	Property value	Deposit size	Monthly interest rate	weekly mortgage(£)	Mortgage term	Additional assumption
Scenario A	Cash buyers	C1	£170,000[33]	-	-	-	-	Household has no preference for any specific type of house or any specific LA in England. Household has no extra money putting on the new home purchase, but have enough money to pay for travel, home moving and other related costs during housing transaction such as stamp duty.
Scenario B	First-time mortgage buyers	F1	-	5%	4.24%	130.38	25	Household has no preference on any specific type of house or any specific LA in England. Interest rate for the whole repayment period are keep the same, here we assume equal mortgage payments. Household has no extra money putting on the new home purchase, but have enough money to pay for travel, home moving and other related costs during housing transaction such as stamp duty.
		F2	-	10%	4.24%	130.38	25	
		F3	-	25%	4.24%	130.38	25	
	Non-first-time mortgage buyers in 2009	NF1	-	5%	4.24%	161.54	25	
		NF2	-	10%	4.24%	161.54	25	
		NF3	-	25%	4.24%	161.54	25	
Scenario C	Londoners who are looking for extra-space to live	HM1	£250,000[34]	-	-	-	-	Household has no preference on any specific type of house or any specific LA in England. Household has no extra money putting on the new home purchase, but have enough money to pay for travel, home moving and other related costs during housing transaction such as stamp duty.

[33] This value is determined by median TP in England in 2009.
[34] This value is determined by median TP in London in 2009

6.3.2.2 New proxy of house affordability at LA level

Here we use the estimated affordable property size (m^2) as a proxy of housing affordability by property type for the typical household buyers shown in Table 6.6. This house affordability by LA is calculated using equation 6.1 below:

$$HA_{pjt} = pv_t/(\beta_0 + l_{0j}) \qquad (6.1)$$

Here HA_{pjt} is the affordable property size for the property type p in LA j in a given year t. There are four property types used in this research, which are detached, semi-detached, terraced and flats/maisonettes. pv_t is the hypothetical property value in the same year t, it also refers the housing budget in this research. β_0 and l_{0j} are parameters from the growth curve equations in Table 6.2 for the property type p in a given year t.

In this new housing affordability index (HA_{pjt}), the larger the value, the more affordable the property for the typical household. If HA_{pjt} is lower than the minimum space standards for new homes in in England it is treated as unaffordable. For the minimum space standards, we use the latest minimum space standards for a single occupier published on 30/9/2020, which is 37 m^2 of floorspace for a new one-bed flat (MHCLG, 2020). 37 m^2 is also the minimum floor area for housing standards for a one bedroom one person dwelling announced in 27/3/2015 (MHCLG, 2015). Thus, 37 m^2 is chosen as the minimum space standard for new homes to assist in our definition of affordability in the housing affordability index (HA_{pjt}).

For the hypothetical property value for scenarios A and C we directly use the dwelling price in Table 6.6. For scenario B, the hypothetical property value is derived from the standard mortgage repayment formula (Levina et al., 2019) as in equation 6.2 below:

pv_t = pm[((1+r)^t -1)/(1+r) ^t *r]/(1-d)= pm[((1+r)^t -1)/(r (1-d)(1+r) ^t] (6.2)

Here pv_t is property value in the year t, pm is the monthly payment. The monthly payment is quantified as the 4.33 times the weekly mortgage payment; d is the deposit size (i.e. 0.05 or 0.10 or 0.25). r is the monthly interest rate and t is the mortgage term (i.e. 300 months).

Due to SL (special license) issue in EHS, the weekly mortgage payment after 2014 is unavailable due the pandemic. Therefore, buyer's year after EHS 2014 is unavailable when this research is conducted. The sample size for buyers in 2009 with weekly mortgage payment information is greater than that for buyers after 2009. Considering this data limitation, the following housing affordability analyses in this chapter are mainly based on 2009 housing buyers in EHS. To further explore the spatial-temporal patterns of housing affordability by property type, a separate analysis based on the third scenario is conducted to illustrate the housing affordability change by property types after one year. The LAs' HPMs in 2010 are estimated through the three-level GCMs in Table 6.2. To simplify the analysis, we assume that this typical household maintains the same wealth and property value as 2009 during these two years. In this approach, the result will offer a directly picture on how the affordability size change for this household if they buy the property one year later.

6.3.3 Results

6.3.3.1 LA housing affordability in 2009 for three scenarios

Based on Table 6.6 and equation 6.2, the property value for each typical household along with the related basic mortgage information is estimated (Table 6.7). The three typical first-time buyers in scenario B who monthly pay £564.55 for 25 years as determined by the 4.24% monthly interest rate, can get £95,730 from the bank. At the end of the 25 years, these three typical buyers will pay an extra 77% of the loan amount to the bank due to the interest rates. This extra pay for the interest rate is the exactly the same for the other three typical non-first-time buyers. This more money a household borrowed from the bank, the more money they will pay in interest by the end of payment term. Looking at the ratio of mortgage interest payments to property value, it will be seen that higher deposit value represents a relative low proportion. For the household who chooses the 5% deposit size, they will pay 73% of the total property value in interest by the end of the payment term. This proportion decreases to 69 % if the deposit size is 10%. It further decreases to 58 % if the deposit size is 25%. Detailed statistics for households in scenario B are in Appendix C5. Although the mortgage approach helps the buyers to get the property at the starting point, they will pay a big proportion to the bank at the end of payment term. This also reveals that reducing the interest rate for the same repayment period year will increase buyer's ability to purchase an expensive property. A mortgage could help buyers to buy a property, but it also reduces their ability to purchase for a larger property.

Table 6.7 The estimation of property value for the three candidate scenarios

Scenario ID	Buyer type	Typical household ID	Property value (£)	Deposit value (£)	Monthly mortgage (£)	Loan amount (£)	Mortgage interest payment (£)
Scenario A	Cash buyers	C1	170,000.00	-			
Scenario B	First-time mortgage buyers in 2009	F1	100,768.10	5,038.41	564.55	95,729.70	73,633.92
		F2	106,366.30	10,636.63	564.55	95,729.67	73,633.92
		F3	127,639.60	31,909.90	564.55	95,729.70	73,633.92
	Non-first-time mortgage buyers in 2009	NF1	124,851.00	6,242.55	699.47	118,608.45	91,231.97
		NF2	131,787.20	13,178.72	699.47	118,608.48	91,231.97
		NF3	158,144.70	39,536.16	699.47	118,608.54	91,231.97
Scenario C	Londoners who are looking for extra-space to live	HM1	250,000.00	-			

(1) Housing affordability in 2009 for scenario A at LA by property types

Using equation 1 and the typical household C1 in Table 6.7, the new housing affordability (affordable property size) by different property type is estimated at LA level in 2009. Figures 6.15 to 6.18 present typical household C1's affordable size at LA level by four property types. LAs shaded in red are unaffordable areas. LAs shaded in yellow represent those locations with less housing affordability (affordable property size) for the typical household C1, LAs shaded in blue represent those locations with more housing affordability for the typical household C1.

Looking at unaffordable LAs by different property types shows they are located in London. Kensington and Chelsea, Westminster, City of London, Camden and Hammersmith and Fulham. These areas spread to Islington and Richmond upon Thames for terraced properties type. The unaffordable LAs further spread to Wandsworth for semi-detached houses and to Tower Hamlets for detached houses.

Looking at the affordable size among the four property types in England at LA level, detached property is the least affordable type. Flats/Maisonettes are the most affordable type at LA level. However, this pattern is not reflected in some LAs. 53% of LAs show detached property type as the least affordable property type and terraced as the most affordable. 43% of LAs show detached property as the least affordable property type and flats/maisonettes as the most affordable. Meanwhile, there are 6 LAs (Tower Hamlets in London, Eastleigh in the South East, Maldon in East of England, South Hams, Poole and Bournemouth in the South West) showing detached property as the least affordable property type and semi-detached as the most affordable property. Although flats/maisonettes are the most affordable property type in Hackney, semi-detached property is the least affordable property type. The City of London is a particular case where only two property types (terraced, flats/maisonettes) are sold, with terraced a bit less affordable than flats/maisonettes.

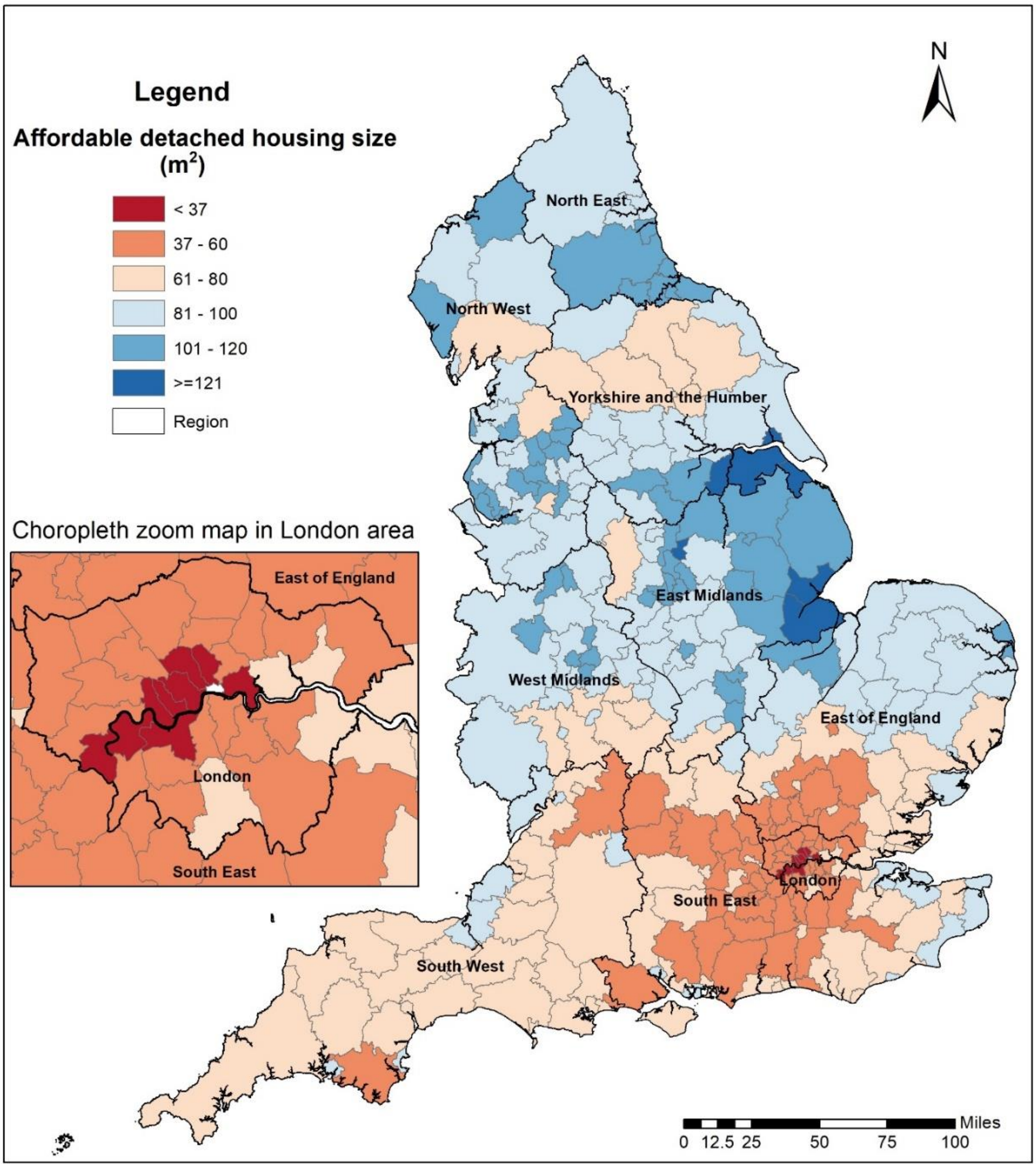


Figure 6.15 The geography of affordable detached property size for typical household C1 at LA level

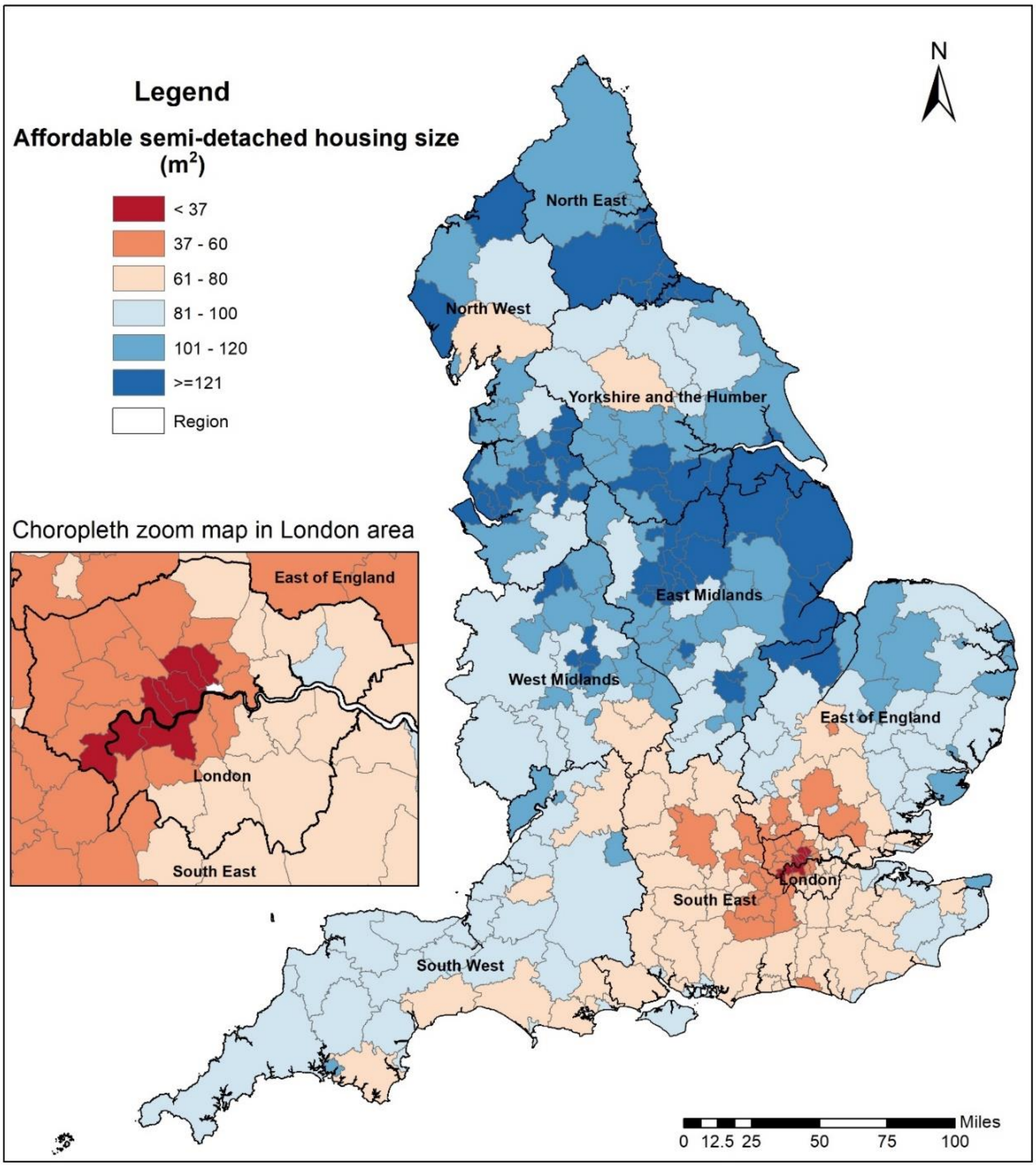


Figure 6.16 The geography of affordable semi-detached property size for typical household C1 at LA level

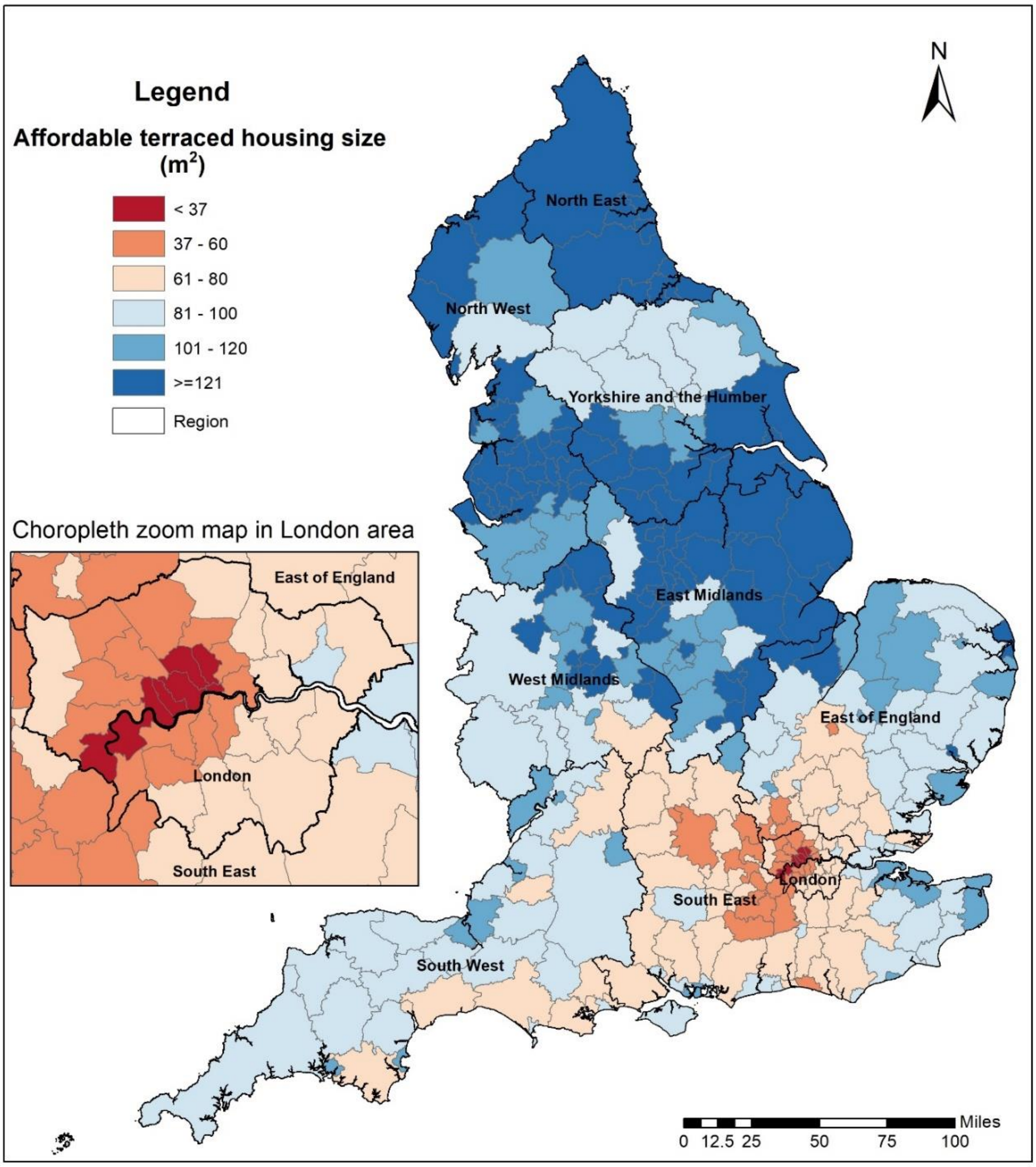


Figure 6.17 The geography of affordable terrace property size for typical household C1 at LA level

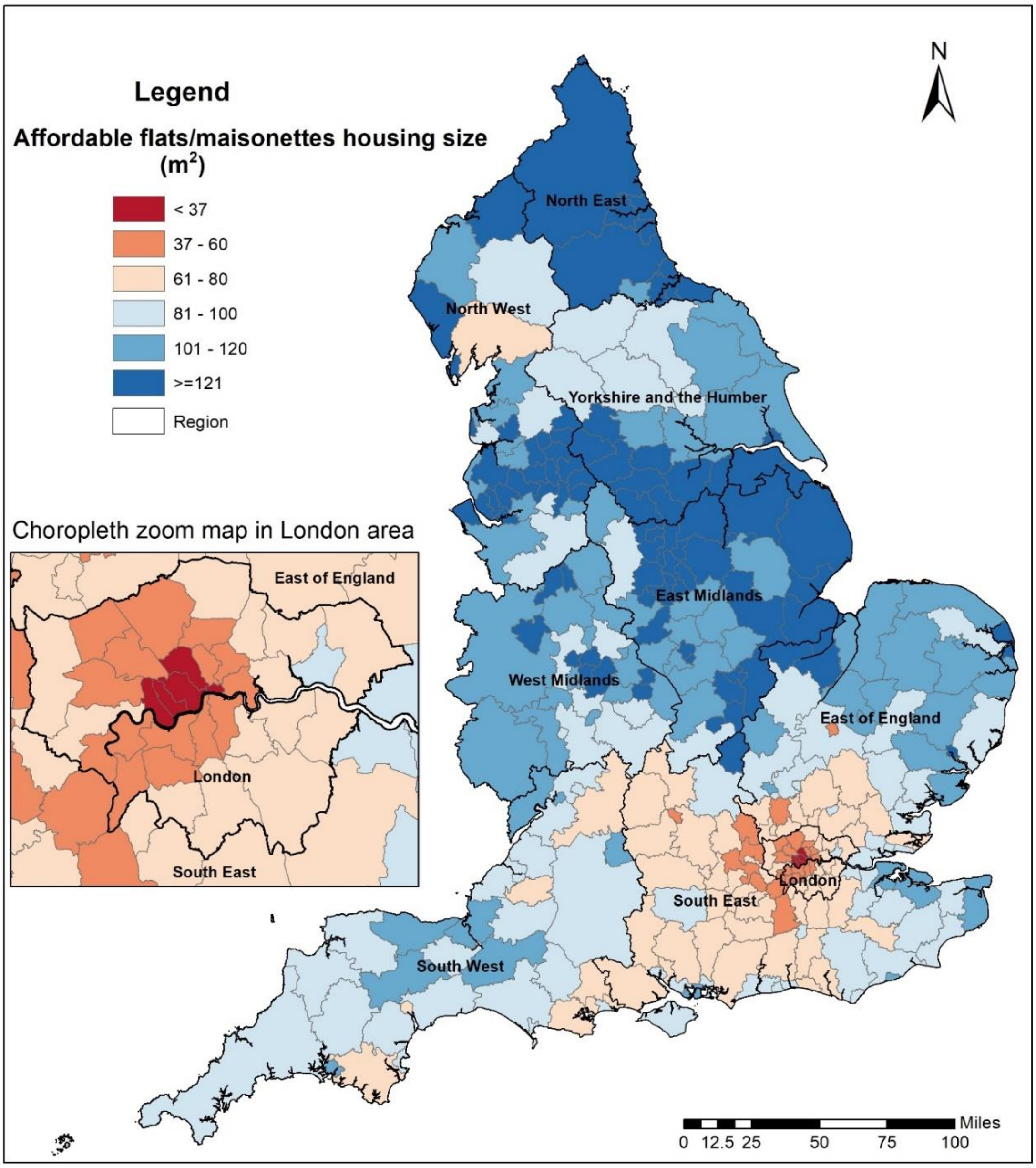


Figure 6.18 The geography of affordable flats/maisonettes property size for typical household C1 at LA level

Looking at the LAs in London, the spatial pattern for affordable size by different property types are roughly the same. LAs located in the west of London around the Thames river are the least affordable area. The least affordable areas spread out to the northwest and south west directions. Other LAs far away the above least affordable areas are become more affordable. It obvious that the East of London is relatively more affordable than the West of London for the same property type. The most affordable

LA within London is Barking and Dagenham.

Table C6.1 in Appendix C6 summarises the rank order relationship of affordable size among different property types for each LA. The affordable size for any two property types within 4 m^2 size difference are treated as approximately equal. According to the most affordable property type, the rank relationship can be crudely sorted into three big groups according to the most affordable property type (Figure C6 in Appendix C6). The most affordable property type for the LAs in group 1 is the terraced house. The most affordable property type for the LAs in group 2 are Flats/maisonettes and terraced houses. The most affordable property type for the LAs in group 3 is Flats/maisonettes. These three groups suggest a North South divide across England. The majority of the LAs in group 1 are located in the North of England with some LAs located near the seaside in South West and South east and East of the England. LAs in group 2 and group 3 are mainly located in the South of England but with a few LAs near Newcastle or in the Yorkshire Dales National Park or the North York Moors National Park. Comparing with Figure 5.11 in Chapter 5, they are the areas which have an HPM over 4,000 £/m^2.

Each group in Figure C6, contains secondary subgroups based on the rank order of affordable property size among the four property types. Details are shown in Table C6.2 in Appendix C6. Table 6.8 summarises the LA count for each secondary subgroup in Table C6.2. Looking at the rank order of the affordable size by LA across England. 125 LAs are classified into group 1 under 7 sub-groups, representing 38.34% of the LAs in England. 74 of these LAs show the most common descending order of affordable size as terrace, flats/maisonettes, semi-detached, detached. The other 26 LAs show a slightly different order in that semi-detached are more affordable than flats/maisonettes, but terraces remain the most affordable and detached the least affordable. The remaining 25 LAs show the same pattern of terraces being the most affordable and detached the least affordable, however, there are variations in the order of the remaining two house types (semi-detached and flats/maisonettes). 18 of these 25 LAs shows the affordability for the middle two property types (the flats/maisonettes and semi-detached) as being equal. The typical city LAs in this group are Manchester, Sheffield and Leeds. 4 of these 25 LAs (Weymouth and Portland, Derbyshire Dales, Cornwall and Hambleton) occupy the same sub-group and show terraced and semi-detached are equally the most affordable type, followed by flats/maisonettes and then detached. Castle Point in East of England comprises a separate group, showing the affordable size order as terrace, flats/maisonettes, equal third order for detached and semi-detached. South Holland in the East Midlands is in a separate group showing the affordable size order as terrace, semi-detached, equal third order for detached and flats/maisonettes. The Isles of Scilly also comprises a separate group, showing terrace and semi-detached are more affordable than flats/maisonettes and detached.

Table 6.8 A description statistic of the LA's affordable size order among four property types

Group No.	Secondary category	LA Count
	Terraced > flats/maisonettes > semi-detached > detached	74
	Terraced > semi-detached > flats/maisonettes > detached	26
	Terraced > flats/maisonettes ≈ semi-detached > detached	18
	Terraced ≈ semi-detached > flats/maisonettes > detached	4
	Terraced > flats/maisonettes > semi-detached ≈ detached	1
	Terraced > semi-detached > flats/maisonettes ≈ detached	1
Group 1	Terraced ≈ semi-detached > flats/maisonettes ≈ detached	1
	Flats/maisonette ≈ terraced ≈ semi-detached > detached	65
	Flats/maisonette ≈ terraced > semi-detached > detached	49
	Flats/maisonettes ≈ terraced ≈ semi-detached ≈ detached	4
	Flats/maisonettes ≈ terraced > semi-detached ≈ detached	2
Group 2	Flats/maisonettes ≈ terraced	1
Group 3	Flats/maisonettes > terraced ≈ semi-detached > detached	53

Group No.	Secondary category	LA Count
	Flats/maisonettes > terraced > semi-detached > detached	18
	Flats/maisonettes > terraced ≈ semi-detached ≈ detached	9

121 LAs are identified in group 2 with 5 sub-groups, which in total represent 37.12% of the LAs in England. LAs in group 2 are the areas that show both terraced and flats/maisonettes are the most affordable property type. Over half of the LAs (65) in group 2 show semi-detached, terraced and flats/maisonettes are similarly affordable and more affordable than detached. There are another 49 LAs showing terraced and flats/maisonettes as the most affordable property types followed by semi-detached and then detached. Four LAs (Canterbury, Islington, Newham, Rochford) occupy the same sub-group and show the affordability for the four property types (flats/maisonettes, terraced, semi-detached and detached) as equal. Slough and Fareham comprise a separate group, showing flats/maisonettes and terraced are more affordable types than semi-detached and detached. London is in a separate group with only two property types, showing affordability for flats/maisonettes and terraced as equal.

LAs in group 3 are mainly located in the South of England with three LAs in the North of England (North Tyneside, Harrogate, and Ryedale). There are 80 LAs in group 2, occupying 24.54% of the LAs in England. of these 53 show the most common descending order of affordable size as flats/maisonettes, equal third order for terraced and semi-detached, then detached. The other 18 shows a descending order of affordable size as flats/maisonettes, terraced, semi-detached and detached. The remaining nine LAs (Broadland, Hackney, Camden, Hammersmith and Fulham, Kensington and Chelsea, Kingston upon Thames, Brighton and Hove, Westminster and Richmond upon Thames) show that flats/maisonettes are more affordable than the rest of the three property types. In these nine LAs, the affordable size for the terraced, semi-detached and detached property types are the same.

For the typical household C1, the affordable size by property type is different at LA scale. Terraced house is the most affordable type for the cash buyers in group 1 LAs, while it changes to flats/maisonettes as the most affordable type for the LAs in group 3. Terraced and flats/maisonettes are equally the most affordable for the LAs in group 2. Within each group, the same buyers face slightly different affordability orders among the remaining three property types. In summary, affordability in terms of property types change at LA level with more complexity in different local LAs.

(2) Housing affordability in 2009 for scenario B at LA by property types

In scenario B, there are six typical households formed from the three different deposit sizes and the two different monthly mortgages payment levels for first time buyers or non-first-time buyers (Table 6.7). None of these six typical households represents a higher housing affordability than the typical cash buyers in scenario A. This is due to the property value for these six typical households being lower than £170,000. This further results in more LAs in England being unaffordable (i.e. affordable property size below than 37 m^2). In equation 6.1, the spatial pattern of affordable property size mirrors the pattern of HPM when controlled for the buyer's budget and property type. For any given buyers in Table 6.9, the spatial pattern of the affordable property size for a given property type exactly follow the spatial pattern of the related HPM. Instead of exploring the affordable property size pattern for the six typical households in scenario B, this section focusses on further exploration of the relationship between the household budget and number of unaffordable LAs.

Table 6.9 A summary of the unaffordable LAs for buyers in scenario B

Buyer type	**Typical household ID**	**Property value (£)**	**Unaffordable LAs count**			
			Detached	Semi-detached	Terraced	Flats/ maisonettes
First-time mortgage buyers in 2009	F1	100,768.10	78	46	39	32
	F2	106,366.30	66	36	32	22
	F3	127,639.60	31	17	13	9
Non-first-time mortgage buyers in 2009	NF1	124,851	34	18	14	10
	NF2	131,787.20	26	15	11	9
	NF3	158,144.70	12	8	8	6

Table 6.9 shows the statistics of unaffordable LAs by property type for the six typical household's scenario B. These six typical households hold different housing budgets (property value) ranging from £100,000 to £160,000. Households F1 and F2 hold a housing budget of around £100,000. Households NF1 and F3 hold quite similar housing budgets of around £125,000. Household NF2 holds a housing budget of around £130,000 and the last Household NF3 holds the greatest housing budget of almost £160.000. It is obvious that a predominance of buyers with the lowest budget (household F1) give rise to more unaffordable LAs. The majority of these unaffordable LAs are in London or nearby London.

For any given household in Table 6.9, the housing affordability changes by property type. The number of LAs with unaffordable detached housing is the largest category, followed by semi-detached, terraced house and flats/maisonettes categories. This means those households with housing budgets below £160,000, have most affordability issues when they seek to purchase a detached house at LA level in

England. This unaffordability situation will ease greatly if they choose to buy a semi-detached house. However, the unaffordability situation will not ease so much if they choose a terraced rather than a semi-detached house. Choosing flats/maisonettes rather than the above three types of house will allow the household find more LAs within which to buy a standard one bedroom size property. The number of unaffordable LAs for flats/maisonettes drops by almost half in comparison to the number of unaffordable LAs for detached.

Looking at the change in the unaffordable LAs count for detached houses together with the housing budgets in the six households, the number of unaffordable LAs drops by half when the housing budget increases from £100,000 to £124,000. There is further drop by a half once the budget increases to around £158,000. The change situation is similar for the other three property types but with a greater degree of increase in the number of unaffordable areas. This confirms that households with lower budgets face relatively higher numbers of unaffordable areas. Given the above, the households with relatively lower housing budgets face worse housing affordability issues than the other household categories and need more support from society, through policy instruments such as “Help to Buy”.

Mortgage buyers with different weekly mortgage payments but the same underlying repayment period and interest rates reveal different patterns of unaffordable LAs in England, with further differences by property type. The weekly mortgage payment for a typical first-time mortgage household is £ 130.38, while the weekly mortgage for typical non-first-time mortgage buyers is £161.54. This difference of £31.16 in weekly mortgage payments with the same interest rate for 25-year repayment period results in more than a doubling in the number of LAs that are unaffordable in England. The unaffordable LAs decrease relatively less for the same mortgage buyers with higher deposit size. Within the same buyer type, the only difference across the three typical buyers is the deposit size. Household F2 has the twice of deposit size of Household F1. Household F3 has the five times the deposit size of Household F1. However, the number of unaffordable LAs decreases but by a relatively lower amount. For example, for an F1 household buying a terraced house there are 39 unaffordable LAs. Once the deposit size changes from 5% to 25%, the number of unaffordable LAs only decreases by a further 33%. This shows that the mortgage buyers with low deposit face more difficulty in buying a property and need effective support to increase affordability, such as a relatively low interest rate.

(3) Housing affordability in 2009 for scenario C at LA level by property types

In scenario C, the typical household HM1 is a London home-owner with a property value of £250,000 looking for a bigger house with the same property value. This typical household HM1 has the highest affordability in comparison with the other typical households in scenario A and B. Figures C7.1 to C7.4 in Appendix C7 illustrate the housing affordability for this typical household HM1. The spatial patterns for affordable property size are similar between property types, with high house prices located in or nearby London, but they vary in some details. Analysis comparing the spatial patterns for affordable property size and each LA's commuting time to London (Figure 6.19), reveals a different spatial pattern. The difference between these two spatial patterns reveals potential opportunities for homeowners to find larger properties with the same commuting time to London.

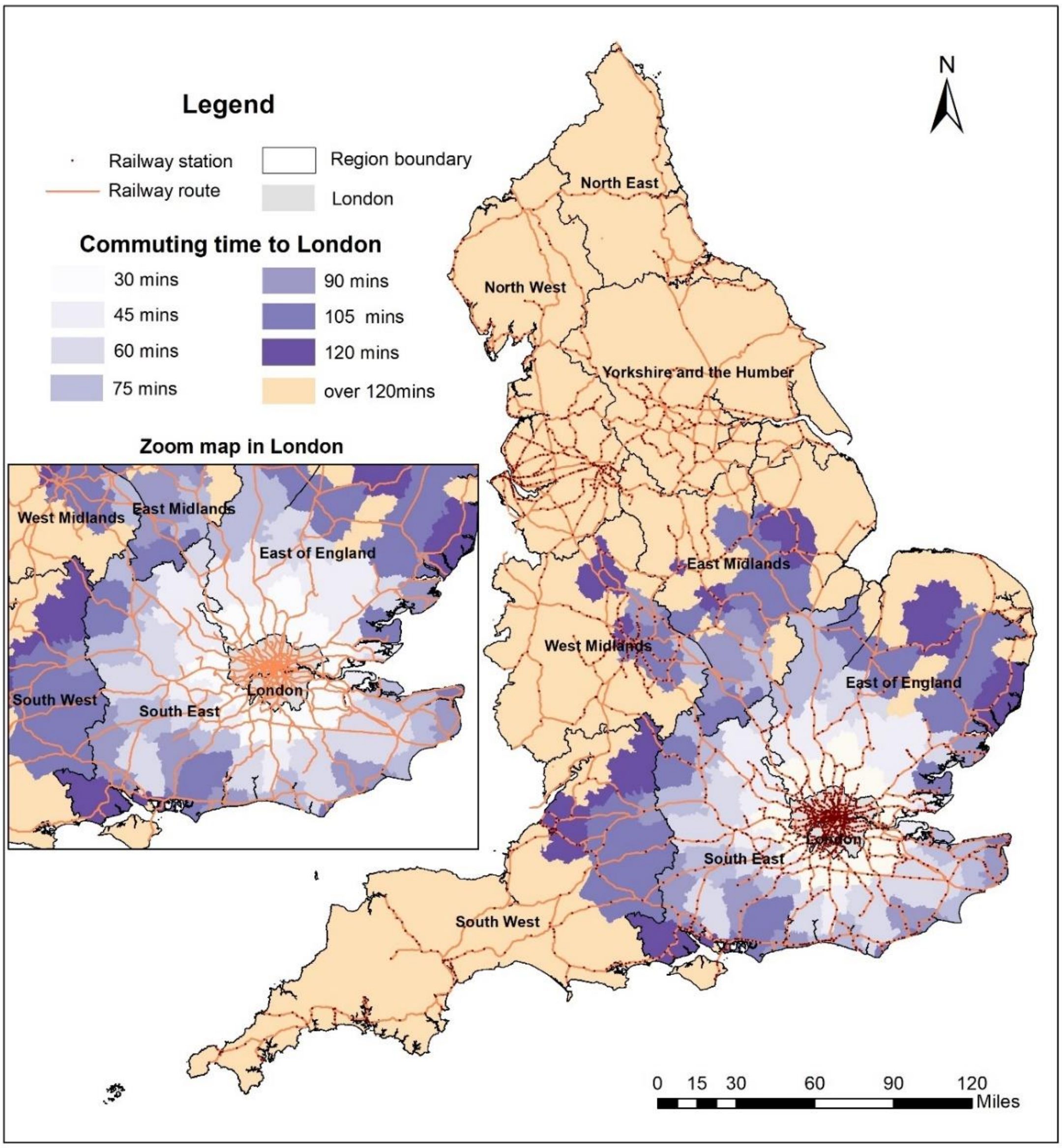


Figure 6.19 The geography of commuting time to London by public transport at LA level

To better understand the variation of affordable property size within the commuting time categories, Figure 6.20 represents the relationship between maximum property size among the four property types and commuting time to London. The most affordable LAs in each commuting time category (below two hours) are labelled. These are the places where the maximum property size can be achieved whilst retaining the same commuting time. Figure 6.21 maps these most affordable LAs (in red boundary) along with their commuting time. The labelled LAs are mainly clustered in two directions centred on London; one is in the northwest direction and more loosely clustered, the other is due east and tightly grouped. Furthermore, the top five most affordable LAs for each commuting time group are located in the same two directions plus a north-eastern direction (Figure 6.22). The northwest direction clearly follows three railway routes: the first from London to Birmingham (West Coast Route), the second from London to Leicester (central route) and the last from London to Peterborough (East Coast Route). The eastern direction follows the railway route from London to Ashford. The northeast direction follows the railway route from London to Ipswich. Thus, the above three directions (i.e. northwest, due east and north-eastern directions) appear to offer the best opportunities to homeowners who are looking for more living space.

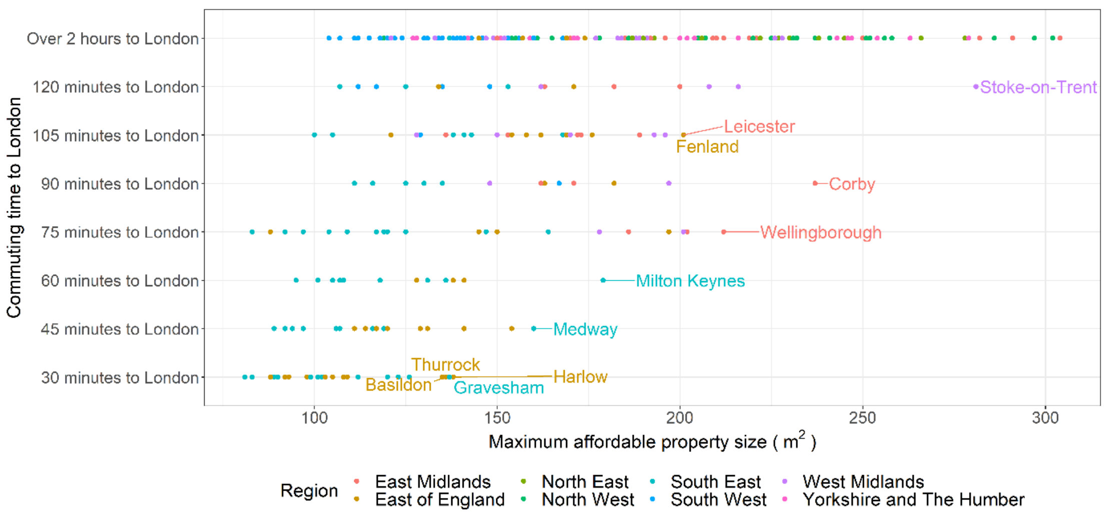


Figure 6.20 The relationship between maximum affordable property size and commuting time to London for the typical homeowner

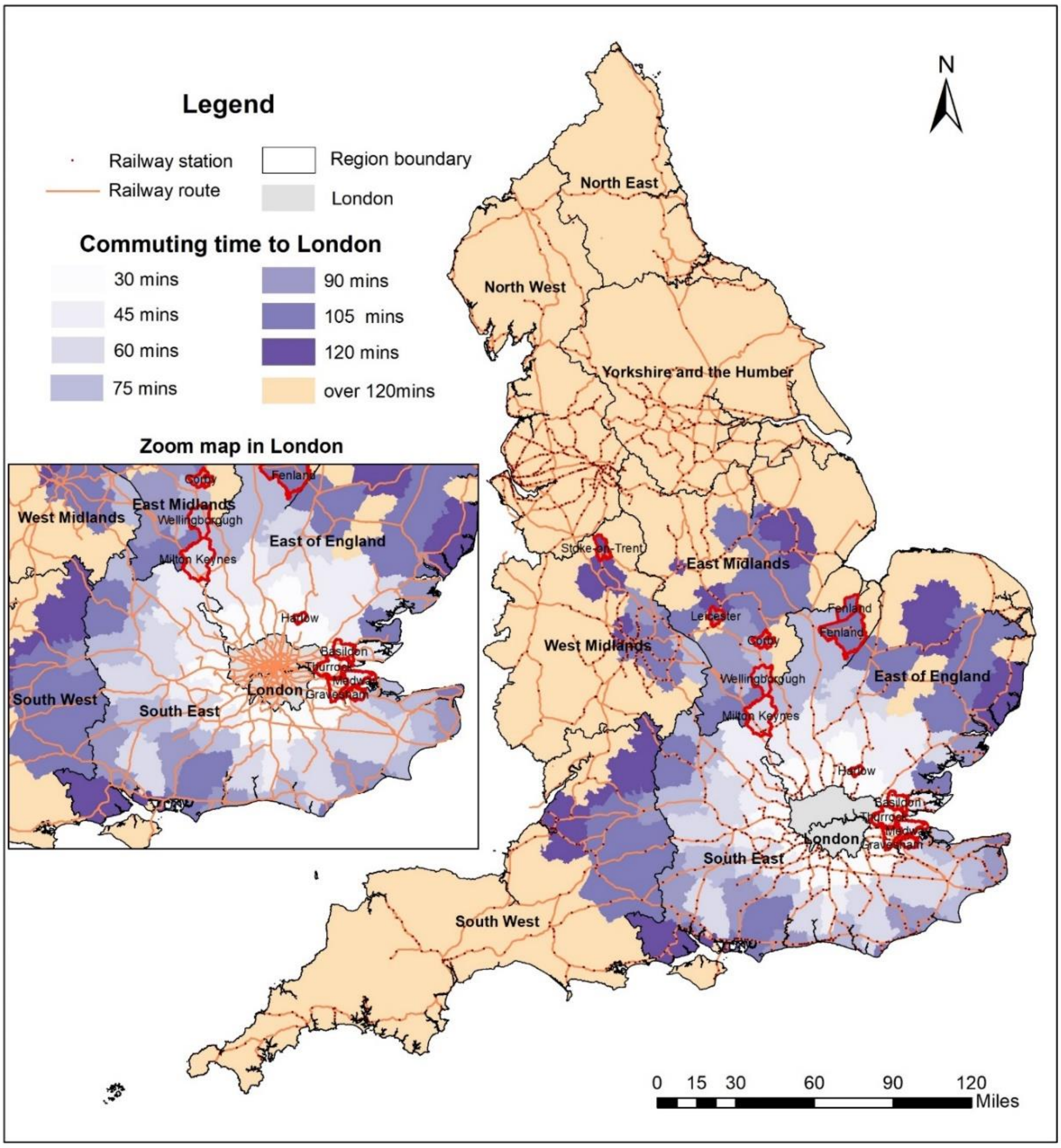


Figure 6.21 The most affordable LAs in England for typical household HM1 moving out of London

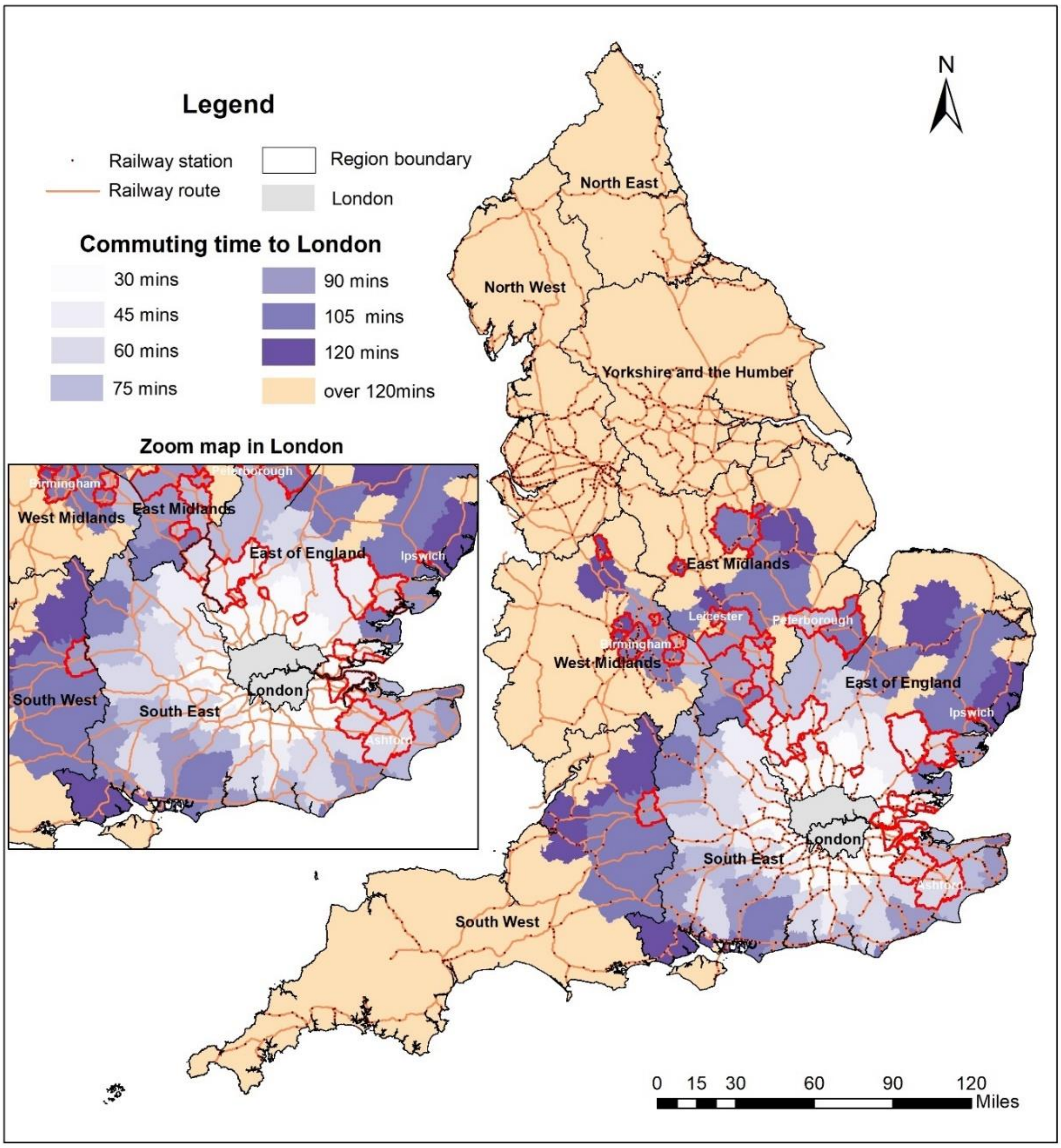


Figure 6.22 The top five most affordable LAs in England for typical household HM1 moving out of London

6.3.3.2 LA housing affordability change between 2009 and 2010 for scenario C

With an understanding of housing affordability for typical cash buyers, mortgage buyers and London home movers, this section aims to further investigate the housing affordable change at LA level. Considering the LA's HPMs are differ by property type and also exhibit different percentage changes, here we explore how these HPM variances shape the affordable size change for given typical buyers. Based on equation 1, the change of affordable property size for a given property type is estimated between

2009 and 2010 (Table C8.1 in Appendix C8). The results show that the change of affordable size for four types ranging from -12.25 m^2 to 7.50 m^2. The change of the affordable size is quite small. According to the 7.5m^2 min floor size for a single bedroom and 11.5m^2 for a double (or twin bedroom) in the nationally described space standard (MHCLG, 2015), this annual LAs affordable floor size change is further classified based on above two standard bedroom sizes. Table C8.2 in Appendix C8 summarises the LAs with affordable size change over a standard single bedroom size. For typical buyers in Corby in the East Midlands who purchased a detached house, terraced house or flat/maisonette one year later than 2009, the buyer loses property size equivalent to a standard single bedroom due to the local HPM change. The affordable size in Corby decreased more for the buyer who purchased a terraced house, the results showing a lose equivalent to a double bedroom. This reveals that buyers with the same budget lose the most property size in Corby if they buy the house one year later. Hartlepool in North East represents the opposite situation of affordable size change for the same buyer who buys a flat/maisonette. Here the buyers will gain one single bedroom size in a flat/maisonette in Hartlepool if they buy the property one year later. This is mainly due to the local flats/maisonettes' HPM showing a decrease between 2009 and 2016. This is the similar to the situation of buying a terrace house in Hartlepool. The buyers with the same budget could gain almost a single bedroom size (7.23 m^2) if they brought the house in 2010 rather than 2009. However, the change of affordable size for detached and semi-detached houses in Hartlepool is relatively small, at around 1 m^2. Similar, County Durham and Sunderland in the North East and Barrow-in-Furness in the North West also represent over 6 m^2 property size increases in terraced houses between 2009 and 2010. These results indicate that the affordable size for the buyers with the same budget could increase if they brought terrace houses or flats/maisonettes in a few LAs in northern England.

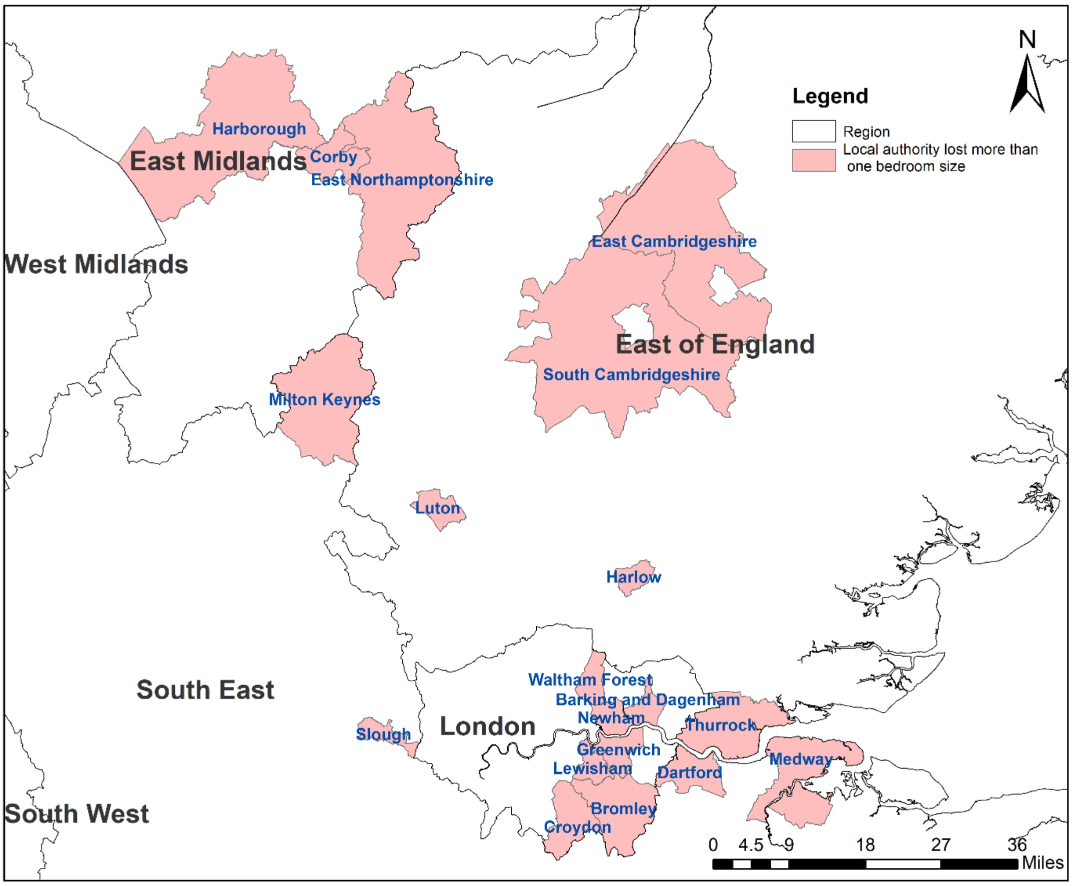


Figure 6.23 The geography of LA with more than one single bedroom size change in affordable size (flats/maisonettes)

The majority of the LAs in England show a less than a single bedroom size (7.5 m^2) decease when the property is purchased in 2010 rather than 2009. There are still quite a few LAs showing a relatively large affordable size decrease. For example, for the buyers who buy flats/maisonettes one year later than 2009, 19 LAs in England lose more than one single bedroom size. Figure 6.23 shows the location of these 19 LAs. Most of these LAs are located in the east of London or the nearby LAs adjoining East London. Compared with the starting-price and HPM percentage change in Figure 6.7 and 6.12, these LAs are the places with a lower HPM (e.g. lower than 3000 £/m^2), which are not the most expensive areas. But their HPMs shows a high percentage increase. This finding reveals that some LAs in England with low HPM but high HPM increase will face more serious affordability issues for the same budget buyers. These findings apply are not only to buyers of flats/maisonettes, but also to buying detached, semi-detached and terraced houses.

6.4 Conclusion

This research adopts a novel research framework to explore the spatio-temporal pattern of housing affordability at LA level in England. Unlike the traditional approach of reflecting housing affordability by using the ratio of median house price compared to income, this new approach creates a research framework to investigate the housing affordability in terms of the affordable property type by three designated typical buyers' scenarios (i.e. cash buyers, mortgage buyers and home movers). This simple approach reflects the complexities of housing affordability using the newly created individual level HPM dataset in Chapter 3, something that has not previously been possible due to the absence of a HPM dataset in the England. The new housing affordability index shows advantages over the traditional house price ratio in three main ways: first, this housing affordability index is constructed by further exploring the house price variation at LA level and simplifying the variety of possible household configurations. In this way, the housing affordability issue can be well reflected in terms of the affordable size according to property type, changing buyers' budgets, moving outside of the home LA or affordable change across time. Secondly, the new housing affordability index enables comparisons across space and time. Using affordable size as the housing affordability proxy offers a meaningful explanation of housing affordability for the public. It is also able to use the UK nationally described space standard to quality affordable size change for the public and the policy makers. What is more, this housing affordability proxy directly reflects the affordable difference in any other location in the UK. Last but not least, the new index can be linked back to the underlying influences of the local HPM trend, which offers insights into how the house price trends shape housing affordability for the designated buyers.

This research shows that the LA's HPM differ across property types, further showing differences in HPM trend between 2009 and 2016. These differential HPM trends contribute to buyers' affordable sizes in different ways. For buyers who buy the property in the same year in England, the affordable size for the same property type is different across England. Meanwhile, the London house moving scenario suggests the affordable size change in three moving directions (northwest, due east and north-eastern directions) which could offer a greater property size but retain the same travel time band. Furthermore, some LAs (e.g. Islington) show no difference of affordable size within LA by property type, but some LAs (e.g. Camden) do show a difference. For buyers with the same housing budgets, who choose to buy property one year later, the change of affordable size is diverse. Major LAs show a loss of affordable size, but some LAs shows a tiny increase in the affordable size for some property types. Moreover, some LAs with lower HPM but a high HPM increasing trends show the biggest affordable size drop down when compared to the most expensive LAs. This research reveals that change of housing affordability and the its starting HPM show different stories at LA level across England. This imply a uniform housing policy or planning policy may be relatively inefficient. In England a suitably diverse local housing and planning policy is needed to reduce the pressure of housing affordability on buyer's

This research offers an innovative approach to understanding the housing affordability issue through a further exploration of house price variation. It offers a new approach to understanding housing affordability through use of affordable property type, but there is still some space for improvement on the understanding of housing affordability. Further research may consider other more potential typical types of buyers (e.g. multi-earner households) with more diversely designated buyers to advance and enrich the understanding of housing affordability. Instead of using the scenario based typical household, a typical agent-based simulation approach will use to systematically explore the housing affordability vary by buyer's budget. More details of conditions for designated buyers will further considering. Such as accounting for travel costs in the moving out of London (or other city based) scenario.

Chapter 7 Thesis discussion and final conclusion

7.1 Introduction

The research reported in this thesis is data-driven research and aims to better understand housing affordability in England through in-depth analysis of residential house price variation. In completing this research there have been a number of achievements: a new comprehensive national house price dataset, along with property size, has been created; a set of technical validation and data cleaning processes have been specified. An improved house price measure (HPM), to better reflect house price variation, has been identified and produced. A series of model-based frameworks, to coordinate and integrate analyses of HPM across different spatial and temporal scales, were constructed. With a systematic understanding of spatial-temporal patterns of house price variation and the better house price variation estimator, a new housing affordability index in terms of affordable size, by property type, has been generated. The consistent research framework casts a new light on the housing affordability landscape of the UK.

This chapter concludes the thesis through summarising the main research findings and achievements. Section 2 summarises findings with specifically address the objectives laid out in Chapter 1. Section 3 addresses some of the limitations of this piece of research which may be addressed through a suggested research agenda for the future. Section 4 concludes the remarks of this thesis.

7.2 Summary of findings and relevance to policy guidance

This thesis offers a new approach to advance understanding of both house prices and housing affordability in England for the time period after the GFC. The research addresses seven specific objectives. This section will show how each of the objectives has been addressed.

1. *To investigate substantive literature and the current methodological techniques on house price variance and housing affordability with a more specific focus on the UK context.*

In order to better understand the housing affordability through in-depth analysis of house price variation, Chapter 2 starts with a general background of house price change and the UK housing crisis since 1953. Reviewing house prices and housing policy over the past 60 years, the UK was facing a long-term increase in house prices. The underlying reason for this was the UK's continuous increase in population combined with limited housing supply after 1990. The review then moved to the general concept between the rising housing cost and its contribution to the housing affordability issue. With a main focus on housing affordability, the debate of housing affordability is discussed in depth in this thesis. A review of the house price variation research, suggested that housing affordability is crudely estimated at a number of geographic scales (e.g. region), without considering the nature of house price variation across different geographic scales. In England, housing affordability is normally represented by the ratio of median TP to income with little further empirical analysis.

Following the above review, a clear gap in the current understanding of housing affordability in England emerged, especially in terms of considering house price variation. Thus, the substantive research question aims to answer the question, "To what extent does residential house price vary at small geographic levels in England, and how can we best characterise this variation? Could the analysis of house price variation at small geographic levels, combined with housing budgets for different types of buyer, help advance our understanding of the spatial and temporal patterns of housing affordability?"

In addition to the general literature review in Chapter 1, Chapters 4 to 6 detail a review of the literature and the current methodological techniques to better support each empirical analysis. Chapter 4 investigates literature in the house price variation research area within the UK. It not only recognises that there is incomplete understanding of the spatial extent of the variation in England's housing market but also chooses the most suitable methodology (multilevel models) to address this research gap. Chapter 5 reviews national and regional house price change trajectories research, showing that empirical house price trajectory studies are mainly conducted at regional level. The regional level house price variation is normally conceptualized as a ripple effect pattern starting with high prices in London. There are, however, a few research papers that start to explore the regional house price variation by property age or property type. Chapter 6 reviews the current housing affordability measures along with the issues these raise.

2. *To examine and review house price datasets and income datasets in England from public open datasets and identify the data deficiencies in understanding the house price variation in order to create methods to overcome data deficiencies.*

In order to better reflect housing affordability, Chapter 3 starts with reviewing the available data resources and summarises their deficiencies. For the open access house price datasets, LR PPD is identified as the most comprehensive residential house price data, in terms of understanding the dynamics of the housing market in England. Rather than the other house price datasets, which represent only a sample of the residential housing or rental markets, LR PPD covers all transaction records in England since 1995. However, the data has some limits that hinder the analysis of the house price variation at small geographic scales. One limitation is that the data are not geo-referenced, which precludes exploration of the house price variation at property address level. The second limitation is the data lacks dwelling size information along with other details of the property's characteristics. Two data linkage methods are created to overcome these shortcomings to link it with OS MMTL, OS ABP and Domestic EPCs. A series of data pre-processing methods containing match rate and two statistical tests (Kolmogorov–Smirnov test and Chi-squared test) was set up. Based on the amount of lost information through the above-mentioned pre-processing, the time period between 2009 and 2016 were identified as the most appropriate period for this research.

Finally, Chapter 3 focussed on reviewing the income data used to explore the housing affordability issues in England. The income data review observed that income data resources are even more limited than the house price data resources. There is no accessible long-term official individual household income data along with detailed household circumstances in the UK. The current available income data are accessible either directly at aggregated level or at survey based individual level. For the survey-based income data, they are unable to detect an individual home address; instead, they only geolocate at macro geographic levels (e.g. by region) so as to protect the personal data. Aggregated income data are first excluded in this research, because it limits analysis at other geographical scales and is unable to reflect differences between households. Within all the observed household income data, the EHS was chosen as the best, in terms of the time coverage and richness of information on household age, income, housing cost (i.e. mortgage payments), and tenure. Considering the data review of house price and household income data, this work selects the enhanced LR PPD from the house price datasets, and EHS data from the income datasets, as the core research datasets for the subsequent investigation. Because the time period for cleaned enhanced LR PPD is between 2009 and 2016, all the following research focuses on this time period.

3. *To build on prior methods and develop a reusable research framework to explore the housing variation at multiple scales and choose an appropriate house price indicator for the given geographical level.*

Following the literature and data review, the empirical research focussed on addressing the research question step-by-step. Chapter 4 starts with exploring the house price variation at multiple geographical scales. Two housing price measures are used for comparison, namely TP and the HPM. Eight pairwise multilevel variance components models are used to estimate variation in two house price measures (i.e. TP and HPM), at four different geographical scales (from LA to individual address), for each year between 2009 and 2016. VPC and ICC for each geographical scale from the multilevel models are used to understand the nature of house price clustering at the observed four spatial scales.

Comparing the VPC for TP and HPM for the same year variance components models, both of these measurements reveal (Figure 4.9) that house price variance does exist between LAs, within-LA-between-MSOAs and also within-MSOA-between LSOAs. HPM variance across different levels follows a similar pattern as TP variance, but with a higher variability at the same level. This reveals that HPM aggregated at geographic level (i.e. LA level or MSOA level or LSOA level) represents more house price variation, and in return offers a more accurate picture of the England's housing market, than TP. The ICC results for the HPM models for each year between 2009 and 2016, show that HPM clusters at the LA level, but there is also an increased degree of clustering at MSOA level. The cluster degree shows a very small increase when moving to the lower geographical level (LSOA level). Overall, HPM variability in England shows an increase from 2009 to 2016 at LA level. In 2009, 53% of house price variation existed between LAs. The magnitude of disparities increased 75% (1.42 times) in the following eight years.

In summary, accounting for the size of properties by using HPM offers a more accurate picture of house price variation than does the use of TPs at the same geographic scale. HPM in England are found to be clustered at LA level and highly clustered at MSOA level between 2009 and 2016. Overall, 50% (+20%) of house price variation is observed at LA level. This is therefore the most appropriate geographical scale at which to begin to understand house price variation in the England. There is an additional 10% of total house price variation observed between MSOAs within their LAs. This indicates that the MSOA is the second most appropriate geographical scale to consider when exploring the house price variation in England. There is no need to consider the LSOA level as the HPM within MSOAs is more highly clustered.

4. *To build on the research findings and further explore temporal house price variation*

To fully understand the nature of spatial and temporal variation in house prices in England between 2009 and 2016, Chapter 5 explored the effect of time on house price variation, with a control of spatial scales. Chapter 4 revealed that HPM are clustered at LA level and highly clustered at MSOA level, thus these two spatial scales are retained in Chapter 5. The whole analysis is divided into three parts comprising two types of multilevel model and a cluster analysis.

The first part aims to understand the extent to which space and time influence HPM variation in England. Four-level variance components models are used to investigate the two spatial effects and one temporal effect on house price variance. For the temporal effects three commonly used time slices in analysis house price trend data (quarter, half-year, and year) are separately investigated through a four-level variance components model, with each model exploring one time effect along with two spatial effects (LA and MSOA). Results reveal that the LA effect contributes 59% of total HPM variance over the time period (2009-2016), with the MSOA effect within the same LA contributing a further 12%. It does not matter which time scale is chosen, the time effect on HPM variance is the same and relatively small (5%). This revealed that HPM within the same MSOA shows a small change across time but HPM between MSOAs within the same LA or between different LAs shows great variation.

A one-year time scale has been found to fit the variance components model best. The second part of this set of analyses starts to use multilevel GCMs to further investigate the house price trend at and below LA level. The model results show the LA house price trends did vary across LAs and overall show a ‘fan out’ growth trend. Moreover, the variation in house price trend between MSOAs, but within LA, is too small to ignore. This reveals that HPM within an LA shows a similar trend.

The third part continues to further unlock the spatial-temporal pattern of HPM variation across LAs in England. The hybrid hierarchical k-means clustering method is used to cluster the spatial pattern of each LA’s estimated HPM and overall HPM trend from the selected GCM. A five-cluster result was chosen as the spatial-temporal house price cluster result at LA level, because the cluster pattern changed only a little when adding one more cluster. Table 5.4 summarises the house price trend for each cluster group. The results indicate that the LA house price spatial-temporal pattern represents five change types, which are: cheap house price area with very little change; moderate house price area with small increase; moderate house price area with moderate increase; high house price area with high increase; very expensive house price area with high increase. Then a spatial map (Figure 5.15) of the clustering result was plotted to delineate the spatial and temporal patterns of the LA level house price in England. The house price spatial-temporal pattern in England at LA level presents a gradient pattern with expensive house price and high increases centred on inner London but decreasing as distance from the centre increases. It offers an LA scale empirical research result demonstrating the London ripple effect which previous research has observed at regional level. What is more, the London influence is not equal in all directions, with a stronger influence in the west direction than in the east.

5. *To consider the spatial and temporal pattern of house price variation and to develop an effective method to reflect spatial-temporal housing affordability for different types of buyers.*

With a clear understanding of LA effect on house price variation in England (Chapter 4) and LA house price trends in England between 2009 and 2016 (Chapter 5). Chapter 6 explores how the LA house price variations shape housing affordability for different buyers. A three-stage workflow was created to offer a novel way to reflect the housing affordability issue. After reviewing all the existing housing affordability measurements in England, the BBC calculator is identified as a more detailed approach to the question of housing affordability, with evident practical value to potential home buyers. But, unlike the latest alternative ONS housing affordability measurement, the BBC approach did not distinguish house price by property type. A new approach has been created by emulating and combining the ideas of the BBC calculator and alternative ONS housing affordability measurements, but based on the better house price indicator identified previously (HPM). Some current housing affordability metrics directly estimate the housing affordability by property type without analysis its variation.

To address this gap, the first stages started with a series of GCMs to explore the spatial-temporal pattern of house price variation in England by four property types (i.e. detached, semi-detached, terraced and flats/maisonettes). This first stage research aims to offer scientific evidence on whether property type needs to be considered at the given level when conducting housing affordability analysis. The results show HPM by property type varies across LA in England, but also shows different trends within the same LA. HPM by property type is used to support the housing affordability estimation. This means house price variation due to property type needs to be considered in the following housing affordability analysis.

By determining the housing affordability estimated from the house price side, Chapter 6 moved on to control for the household budget. The household budget is created based on the local TP or the household housing cost from the EHS. Typical households were designated into three overlapping categories: cash buyers in England, mortgage buyers in England, and London homeowners. The first household scenario centres on cash buyers and aims to explore the LA level affordability change for the most common cash buyer household that buys at the median English TP. This type of household could be represented by mortgage buyers with a 100% deposit. The second scenario focusses on analysing the affordability for one type of mortgage buyer with different levels of deposit. This creates two typical mortgage buyers, with the most common weekly mortgage payment (the median weekly mortgage from EHS) for first-time buyers and non-first-time buyers in a given year, then a further three different deposit sizes for each typical mortgage buyer. The third scenario focusses on home movers to reflect the spatial housing affordability change by commuting distance out of the current home city. London as the most expensive house price location was chosen as an example, aiming to address to what extent could home owners in London with a median value property afford to buy a bigger property by commuting from outside London.

With a clear decision on both house price and household budget considerations, from the first two stages, a new housing affordability index is estimated at LA level in the third stage. Unlike all the existing housing affordability measurements in the UK, the affordability size is calculated to reflect the housing affordability through enhancements derived from the HPM datasets. This new approach is thus able to reflect the degree of housing affordability using the affordable housing size, which is directly meaningful for society and able to identify unaffordability by further considering the standard housing size in England. With this new housing affordability measurement, the affordable size, by property type, for all the three scenarios is explored in detail. Each scenario offers evidence-based information for society to enhance its understanding of housing affordability within a given focus.

6. *To offer specific recommendations on current UK housing policy and planning policy.*

Understanding the nature and extent of differentials in property prices, at different geographical levels, and housing affordability by households with varying budgets, offers a better evidence base for UK housing policy and planning policy. Findings for each empirical chapter (Chapter 4 to 6) bring forth some suggestions for UK housing policy and planning policy. This sub-section will now offer an overall summary.

Chapter 4 revealed that the HPM metric demonstrates more variation than TP at the same geographical scale. Meanwhile, TP will offer a misleading result for total house price variation change (Section 4.4.1). The underlying reason is that TP is influenced by the different mixes of stock bought and sold in different years, or different locations, in total house price variation. Moreover, HPM by property type (detached, semi-detached, terraced and flats/maisonettes) indicates different trends *within* LA areas and further differences *between* LAs (Section 6.2.3). The underlying reason is the differing makeup of property types within each LA across England. All this empirical evidence reveals that the diversity of property size and property type contributes to the complexity of each LA owner-occupied housing market. Different authorities are likely to need different housing policies and planning policies, dependent upon the characteristics of each, based upon the need to recognise the impacts of property size and property type. The delivery of the right type and size of residential housing, across LAs with diverse characteristics, will help ease housing affordability issues. HPM data should be one of the core pieces of information for policy makers or local planners to be aware of and use.

TP does not reflect housing size variation in the local housing market, which may offer misleading quantitative housing market information for policy decisions. Also, to enhance the understanding of house price variation, policy makers or local planners, and even the public are likely to find it worthwhile to look at the HPM at, or below, the LA spatial scale across England. The current regional house price statistics, based on TP data, show weak house price variation information and are less able to reflect the total house price variation and provide as much support for decision makers.

The current housing policy (e.g. Help to Buy or Right to Buy), which differentiates between London and the rest of the UK would be more consistent if it were based on HPM and included some of the more expensive areas bordering London. Housing policy should fit the spatial-temporal pattern of HPM at LA level. Looking at only the spatial patterns of LA starting-price in 2009 (Figure 5.11) and related geography of the overall house price change in the period studied (Figure 5.9), the patterns are not so simple as suggested by the housing policy. Moreover, the geography of spatial-temporal cluster patterns gives rise to the same conclusion (Figure 5.13). Thus, The UK needs a more flexible and locality-based policy, rather than a crude uniform policy. To better support and develop such a policy, more related empirical data-based evidence needs to be gathered and made accessible.

Housing affordability analysis for scenario C (London home mover) reveals three directions appearing to offer the best opportunities to homeowners who are looking for more living space. To ease housing delivery pressure in London, especially in the most expensive central London LA areas, this research suggests a new combined authorities' policy to offer a housing delivery co-operation agreement between the LAs in London and LA areas within a suitable commuting distance of London. For example, two LAs (e.g. hypothetically Camden and Harlow) could set up a housing delivery task agreement to deliver the housing together. Harlow helped Camden Council to achieve its housing delivery tasks by building homes in Harlow. In return, Camden and Harlow Councils use the saved budgets to improve local infrastructure and Harlow's local living environment. Meanwhile, two LA are seeking achievable approaches to shorten the traveling time and costs as further benefits to commuters into Camden. The estimated HPM of detached houses in 2009 is 2,295 £/m^2 in Harlow, while it increases to 7,447 £/m^2 in Camden. This represents a difference of more than 5,000 £/m^2 HPM, which enables Camden to deliver larger and more quality detached houses in Harlow, rather than deliver the same number of small sized detached houses in Camden, for the same cost to the Camden.

The current traditional measure of housing affordability (ratio approach) has limited usefulness when trying to reflect local housing affordability issues, not only due to the macro geographical scales used but also due to its poor representation of housing affordability complexity. UK housing policy and planning policies need to consider or create alternative, more practical, housing affordability metrics to efficiently guide policy development. This research offers a metric to reflect a given buyer's housing affordability in terms of housing size at a given geographical scale, but this should not be the only one. Without a clearer and fuller understanding of housing affordability issues – in terms of where, by whom, to what extent and why these exist – policies aimed at improving housing affordability may fall short of their objectives. Greater engagement by policymakers with academic research, seems likely to enhance the development of measures to help address issues around housing and planning policy, exploiting further potential to deliver richer datasets for policy development. For example, expanding the datasets to include planning consents and changes of use.

7.3 Limitations and future studies

This thesis has advanced both understanding of house price and housing affordability in England with a newly created data set and a series of model-based analyses, but there are some limitations that could be reduced in the future to better understand housing affordability and support the housing policy and planning policy within UK.

The studies built up two complex address-based linkage methods to create geo-referenced spatial house price data in England. After navigating some complex licensing issues relating to the original datasets, one of the linkage methods has been successfully published with the updated linked data in the UK Data Service (Chi et al., 2020), together with a related article published by UCL Open Environment (Chi et al., 2021). This is an address-based linkage between LR-PPD and Domestic EPCs. This published data linkage method shows a similar performance with a slightly lower match rate when updated with new published house prices and also covers Wales. This shows that the existing method has space to improve in the future. for example, the correction for mismatched address strings for England's transactions after 2016 or transactions in Wales. To further benefit society with this reproducible and updatable geo-referenced spatial house price data for academic research, future work will be seeking opportunity to publish the unpublished data linkages (geo-referenced parts with OS ABP) with a proper license.

The research set up a series of model-based research frameworks to investigate the house price variation in England for the time period between 2009 and 2016. To validate the reproducibility and transparency of the research, future study will focus on conducting similar research for the new time and spatial coverages. For example, practical analysis with the later time period (e.g. 2016-2020) will be conducted to examine the research transferability and fully track the understanding of house price and housing affordability patterns after GFC. Furthermore, empirical practices for more wider spatial coverages (e.g. the Greater Britain) are needed to test the analytic replicability and to further hone a series of sustainable methods.

The spatial and temporal pattern of residential house price and housing affordability are well represented at LA level. Nevertheless, considering that an extra 10% variation lies in between MSOAs and within LA, future study will further unlock the house price variation at MSOA level. With a complete understanding of house price variation at and below LA level, society will gain a comprehensive understanding of the dynamic underlying the owner-occupied housing market.

This research explores the spatial and time effects on house price variation in terms of HPM. The underlying driving factors in localised house price variation in different parts of the country are unknown. Having a comprehensive understanding of the main factors driving localised price variations will enable tailoring of effective localised policy responses to housing affordability issues. Future research intends to extend this work through a more thorough exploration of the interacting effects of time, location and key local factors such as plot size, accessibility, the cost of transport, land use structure, housing density and local physical and socio-economic environments. Understanding the underlying mechanisms of house price variation will not only offer deeper insights into pressing housing inequality issues in England, but also a comparable evidence base for housing policy making for the other three countries in the UK.

The housing affordability analysis within this research has been developed with three buyers' scenarios. It provided a new research approach to understanding household affordability variation by different housing budget with some degree of success. However, the current three scenarios only reflect the housing affordability for eight different housing budgets. There is still a knowledge gap in understanding the housing affordability variations within all the other potential housing budgets and how this varies for household structure (e.g. single couple, couple with one child and so on). Therefore, future study will seek a suitable technique (e.g. an agent-based simulation approach) to address this defect. Better understanding the complexity of affordability knowledge in terms of the housing budgets and households' structure will offer better support for housing and planning policy in order to ease housing affordability issues. At the same time, further research will also seek co-operative opportunities with LAs to transfer and validate this research ready for use in their daily governance.

This research offers a new insight into the owner-occupied housing market with a series of analyses to explore the spatial and temporal pattern of house price and housing affordability. However, the understanding of related patterns in the rental housing market and how it shapes the rental housing affordability in terms of property size in different parts of the country is still not well understood. Understanding the dynamic interaction between the rental market and the owner-occupied housing market will be extremely useful to efficiently support housing and planning policies. Thus, future research will consider creating the rental price per square metre with the open access dataset and conduct the same analysis approach to enrich the existing housing affordability research framework with data from the Private Rented Sector.

7.4 Concluding remark

This thesis has explored house price variation in England from 2009 to 2016 and has developed a full overview of spatial and temporal housing price variation at LA level. With a systematic understanding of the house price variation, the thesis provides a new housing affordability index to enable better understanding of housing affordability in England. All these investigations have been achieved based on a newly created HPM dataset. There is, of course, research to be continued, data to be further enhanced and updated, extension to the LA level and to MSOA level, techniques to be improved and patterns for years after 2016 to be observed, but this work has succeeded in unravelling some of the complexities within housing affordability from the affordability size dimension for the first time. It is hoped that through the novel approaches described in this research, those who wish to make similar sense of future patterns of residential house price and housing affordability, both in England and any nation, will now have a new research frame to assist them in achieving those ends.

References

Aha B, Higgins D and Lee T (2018) UK Political Cycle and the Effect on National House Prices: An Exploratory Study. In: *European Real Estate Society 25th Annual Conference*, Reading, 23 March 2018. Available at: http://www.open-access.bcu.ac.uk/6226/ (accessed 4 September 2021).

Ahlfeldt GM, Holman N and Wendland N (2012) *An assessment of the effects of conservation areas on value*. English Heritage. Available at: https://historicengland.org.uk/content/docs/research/assessment-ca-value-pdf.

Ahmed K (2017) Low-income tenants battle soaring rents. 13 October. Available at: http://www.bbc.co.uk/news/business-41601455 (accessed 28 March 2018).

Alakeson V (2011) *Making a Rented House a Home: Housing solutions for 'generation rent'*. Available at: https://www.resolutionfoundation.org/app/uploads/2014/08/Making-a-Rented-House-a-Home.pdf.

Alexander C and Barrow M (1994) Seasonality and Cointegration of Regional House Prices in the UK. *Urban Studies* 31(10): 1667–1689. DOI: 10.1080/00420989420081571.

Andrews J (2020) Massive rise in Londoners moving out of the capital for cheaper homes up north. Available at: https://www.mirror.co.uk/money/massive-rise-londoners-moving-out-21269217 (accessed 5 August 2021).

Ashworth J and Parker SC (1997) Modelling Regional House Prices in the UK. *Scottish Journal of Political Economy* 44(3): 225–246. DOI: 10.1111/1467-9485.00055.

Bailey D, Shah P, Dooley J, et al. (2020) House price calculator: Where can I afford to rent or buy? *BBC News*, 9 July. Available at: https://www.bbc.com/news/business-23234033 (accessed 6 January 2021).

Bank of England (2020) Further details about quoted household interest rates data. Available at: http://www.bankofengland.co.uk/statistics/details/further-details-about-quoted-household-interest-rates-data (accessed 31 January 2021).

Bartholomew DJ, Steele F, Moustaki I, et al. (2008) *Analysis of Multivariate Social Science Data*. Second Edition. Boca Raton: Routledge.

Barton C and Wilson W (2018) What is affordable housing? Available at: http://researchbriefings.parliament.uk/ResearchBriefing/Summary/CBP-7747 (accessed 4 April 2018).

Basu S and Thibodeau TG (1998) Analysis of Spatial Autocorrelation in House Prices. *The Journal of Real Estate Finance and Economics* 17(1): 61–85. DOI:

10.1023/A:1007703229507.

Best R (1996) Successes, failures, and prospects for public housing policy in the United Kingdom. *Housing Policy Debate* 7(3). Routledge: 535–562. DOI: 10.1080/10511482.1996.9521232.

Black RT and III JD (1996) The use of information versus asking price in the real property negotiation process. *Journal of Property Research* 13(4): 287–297. DOI: 10.1080/095999196368808.

Bogdon AS and Can A (1997) Indicators of Local Housing Affordability: Comparative and Spatial Approaches. *Real Estate Economics* 25(1): 43–80. DOI: 10.1111/1540-6229.00707.

Bowlinson L (2019) Why money is pouring into Cardiff's property market. Available at: https://www.telegraph.co.uk/property/house-prices/money-pouring-cardiffs-property-market/ (accessed 14 January 2020).

Boyle MA and Kiel KA (2001) A survey of house price hedonic studies of the impact of environmental externalities. *Journal of real estate literature* 9(2): 117–144.

Bramley G (1994) An affordability crisis in British housing: Dimensions, causes and policy impact. *Housing Studies* 9(1): 103–124. DOI: 10.1080/02673039408720777.

Burke T and Ralston L (2004) Measuring housing affordability. 45: 4.

Card J (2014) Why tech industries are thriving in the south-west of England. *The Guardian*, 10 February. Available at: https://www.theguardian.com/small-business-network/2014/feb/10/bristol-tech-industries (accessed 28 February 2020).

Chandler D and Disney R (2014) *Measuring house prices: a comparison of different indices*. Available at: https://ifs.org.uk/bns/bn146.pdf.

Charlton C, Rasbash J, Jones K, et al. (2019) *MLwiN Version 3.03*. Bristol: Centre for Multilevel Modelling, University of Bristol.

Cheshire P and Sheppard S (1995) On the Price of Land and the Value of Amenities. *Economica* 62(246): 247–267. DOI: 10.2307/2554906.

Chi B, Dennett A, Oléron-Evans T, et al. (2020) A new attribute-linked residential property price dataset for England and Wales 2011-2019. UK Data Service. DOI: 10.5255/UKDA-SN-854240.

Chi B, Dennett A, Morphet R, et al. (2020) Exploring local authority travel time to London effects on spatio-temporal pattern of local authority house prices variation in England. In: *CASA Working Paper 218*. Available at: https://www.ucl.ac.uk/bartlett/casa/publications/2020/apr/casa-working-paper-218 (accessed 7 May 2020).

Chi B, Dennett A, Oléron-Evans T, et al. (2021) A new attribute-linked residential property price dataset for England and Wales, 2011–2019. *UCL Open Environment*. UCL Press. DOI: 10.14324/111.444/ucloe.000019.

Clarke B (2015) The challenge facing first-time buyers. Available at: https://www.cml.org.uk/news/news-and-views/723/ (accessed 26 November 2018).

Collinson P (2014) Data reveals full extent of house affordability crisis in England. *The Guardian*, 23 May. Available at: https://www.theguardian.com/money/2014/may/23/data-reveals-full-extent-house-affordability-crisis-england (accessed 18 November 2018).

Cook S (2003) The Convergence of Regional House Prices in the UK. *Urban Studies* 40(11): 2285–2294. DOI: 10.1080/0042098032000123295.

Cook S (2005) Regional house price behaviour in the UK: application of a joint testing procedure. *Physica A: Statistical Mechanics and its Applications* 345(3): 611–621. DOI: 10.1016/j.physa.2004.07.031.

Cook S (2006) A disaggregated analysis of asymmetrical behaviour in the UK housing market. *Urban Studies* 43(11): 2067–2074.

Cook S and Holly S (2000) Statistical Properties of UK House Prices: An Analysis of Disaggregated Vintages. *Urban Studies* 37(11): 2045–2051. DOI: 10.1080/713707230.

Cook S and Watson D (2016) A new perspective on the ripple effect in the UK housing market: Comovement, cyclical subsamples and alternative indices. *Urban Studies* 53(14): 3048–3062. DOI: 10.1177/0042098015610482.

Cooper C, Orford S, Webster C, et al. (2013) Exploring the Ripple Effect and Spatial Volatility in House Prices in England and Wales: Regressing Interaction Domain Cross-Correlations against Reactive Statistics. *Environment and Planning B Planning and Design* 40(5): 763–782. DOI: 10.1068/b37062.

Cosslett RL (2017) The bank of Mum and Dad: 'It's such a huge amount of money. And guilt'. Available at: http://www.theguardian.com/money/2017/nov/11/generation-rent-property-borrowing-from-mum-and-dad-guilt (accessed 23 March 2018).

Coulter R (2017) Social Disparities in Private Renting Amongst Young Families in England and Wales, 2001-2011. *Housing, Theory and Society* 34(3): 297–322. DOI: 10.1080/14036096.2016.1242511.

Day SB (2016) The 12th Annual Demographia International Housing Affordability Survey. Available at: https://fcpp.org/2016/01/26/the-12th-annual-demographia-international-housing-affordability-survey/ (accessed 18 January 2018).

DCLG (2017) *Fixing our broken housing market*. Available at: https://www.gov.uk/government/publications/fixing-our-broken-housing-market (accessed 4 September 2018).

De Nadai M and Lepri B (2018) The economic value of neighborhoods: Predicting real estate prices from the urban environment. In: *IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*, Italy, 7 August 2018, pp. 323–330. IEEE. Available at: http://arxiv.org/abs/1808.02547 (accessed 20 August 2018).

Demographia International Housing Affordability Survey (2018) *14th Annual Demographia International Housing Affordability Survey*. 22 January. Available at: http://www.demographia.com/ (accessed 14 February 2018).

Dennett A (2010) *Understanding internal migration in Britain at the start of the 21st century*. ProQuest Dissertations Publishing. Available at: http://search.proquest.com/docview/1040511759/?pq-origsite=primo (accessed 7 January 2020).

Department for Communities and Local Government (2017) *Fixing our broken housing market*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/590464/Fixing_our_broken_housing_market_-_print_ready_version.pdf.

Di ZX (2001) *The Role of Housing as a Component of Household Wealth*. Available at: https://www.jchs.harvard.edu/sites/default/files/di_01-6.pdf.

Dong G, Harris R, Jones K, et al. (2015) Multilevel Modelling with Spatial Interaction Effects with Application to an Emerging Land Market in Beijing, China. *PLOS ONE* 10(6): e0130761. DOI: 10.1371/journal.pone.0130761.

Dorling D (2014) *All That Is Solid: How the Great Housing Disaster Defines Our Times, and What We Can Do About It*. Penguin UK.

Doward J (2016) Bank of Mum and Dad is now paying the rent, too. Available at: http://www.theguardian.com/money/2016/jul/16/bank-mum-dad-generation-rent (accessed 24 March 2018).

Downes S (2018) ‘Location, location, location’ still determines house prices. In: *What Investment*. Available at: https://www.whatinvestment.co.uk/location-location-location-is-still-the-most-important-factor-in-house-prices-2614766/ (accessed 2 November 2018).

Drake L (1995) Testing for Convergence between UK Regional House Prices. *Regional Studies* 29(4): 357–366. DOI: 10.1080/00343409512331349023.

Easton M (2013) Rent ‘unaffordable’ for low-income families in third of UK. 15 July. Available at: https://www.bbc.com/news/business-23273448 (accessed 18

November 2018).

Edwards M (2016) The housing crisis and London. *City* 20(2): 222–237. DOI: 10.1080/13604813.2016.1145947.

Everitt BS, Landau S, Leese M, et al. (2011) *Cluster Analysis*. 5th edition. Chichester, West Sussex, U.K: Wiley. Available at: https://www.dawsonera.com/abstract/9780470977804 (accessed 7 January 2020).

Ezennia IS and Hoskara SO (2019) Methodological weaknesses in the measurement approaches and concept of housing affordability used in housing research: A qualitative study. *PLOS ONE* 14(8). Public Library of Science: e0221246. DOI: 10.1371/journal.pone.0221246.

Felton C (2018) Moving out of London - why are people leaving and where are they going? In: *Kiwi Movers*. Available at: https://www.kiwimovers.co.uk/news/moving-london-people-leaving-going/ (accessed 5 August 2021).

Feng Y (2016) *How much is worth? novel quantitative approaches to understanding the changing geography of house prices in England*. Ph.D. University of Bristol. Available at: https://ethos.bl.uk/OrderDetails.do?did=8&uin=uk.bl.ethos.701378 (accessed 18 March 2019).

Feng Y and Jones K (2016) Postcode or Census geography? An examination of neighbourhood classifications for house priced predictions. In: *22nd Annual Pacific Rim Real Estate Society Conference*, 18 January 2016.

Finch WH, Bolin JE and Kelley K (2014) *Multilevel Modeling Using R*. 1 edition. Boca Raton, FL: Routledge.

Fingleton B, Fuerst F and Szumilo N (2019) Housing affordability: Is new local supply the key? *Environment and Planning A: Economy and Space* 51(1). SAGE Publications Ltd: 25–50. DOI: 10.1177/0308518X18798372.

Fotheringham AS, Crespo R and Yao J (2015a) Exploring, modelling and predicting spatiotemporal variations in house prices. *The Annals of Regional Science* 54(2): 417–436. DOI: 10.1007/s00168-015-0660-6.

Fotheringham AS, Crespo R and Yao J (2015b) Geographical and Temporal Weighted Regression (GTWR). *Geographical Analysis* 47(4): 431–452. DOI: 10.1111/gean.12071.

Fotheringham AS, Yang W and Kang W (2017) Multiscale Geographically Weighted Regression (MGWR). *Annals of the American Association of Geographers* 107(6). Taylor & Francis: 1247–1265. DOI: 10.1080/24694452.2017.1352480.

Fuerst F, McAllister PM, Nanda A, et al. (2013) *Is Energy Efficiency Priced in the Housing Market? Some Evidence from the United Kingdom*. ID 2225270, SSRN Scholarly Paper, 26 February. Rochester, NY: Social Science Research Network. Available at: https://papers.ssrn.com/abstract=2225270 (accessed 31 May 2018).

Galster G and Lee KO (2021) Housing affordability: a framing, synthesis of research and policy, and future directions. *International Journal of Urban Sciences* 25(sup1). Routledge: 7–58. DOI: 10.1080/12265934.2020.1713864.

Garcia-Castellanos D and Lombardo U (2007) Poles of inaccessibility: A calculation algorithm for the remotest places on earth. *Scottish Geographical Journal* 123(3): 227–233. DOI: 10.1080/14702540801897809.

Gibb K and Bailey N (2016) Data Scoping Study for a UK Housing Evidence Centre. ESRC report. Available at: https://esrc.ukri.org/files/funding/funding-opportunities/uk-housing/data-scoping-study-for-a-uk-housing-evidence-centre/.

Gibbons S and Machin S (2003) Valuing English primary schools. *Journal of Urban Economics* 53(2): 197–219. DOI: 10.1016/S0094-1190(02)00516-8.

Giussani B and Hadjimatheou G (1991) Modeling Regional House Prices in the United Kingdom. *Papers in Regional Science* 70(2): 201–219. DOI: 10.1111/j.1435-5597.1991.tb01728.x.

Giussani B and Hadjimatheou G (1992) House prices: An econometric model for the U.K. *Netherlands Journal of Housing and the Built Environment* 7(1): 31–58. DOI: 10.1007/BF02496709.

Goldstein H (2010) *Multilevel Statistical Models*. 4th Edition. Wiley Series in Probability and Statistics. Hoboken, NJ, Chichester: Wiley. Available at: https://www.dawsonera.com/guard/protected/dawson.jsp?name=https://shib-idp.ucl.ac.uk/shibboleth&dest=http://www.dawsonera.com/depp/reader/protected/external/AbstractView/S9780470973400 (accessed 4 January 2020).

Goodman AC and Thibodeau TG (1995) Age-Related Heteroskedasticity in Hedonic House Price Equations. *Journal of Housing Research; Washington, D.C.* 6(1): 25–42.

Goodman AC and Thibodeau TG (1998) Housing Market Segmentation. *Journal of Housing Economics* 7(2): 121–143. DOI: 10.1006/jhec.1998.0229.

Goodman AC and Thibodeau TG (2003) Housing market segmentation and hedonic prediction accuracy. *Journal of Housing Economics* 12(3): 181–201. DOI: 10.1016/S1051-1377(03)00031-7.

Gray D (2012) District House Price Movements in England and Wales 1997–2007: An Exploratory Spatial Data Analysis Approach. *Urban Studies* 49(7): 1411–1434.

DOI: 10.1177/0042098011417020.

Gray D (2015) Are prices of New dwellings different? A spectral analysis of UK property vintages. *Cogent Economics & Finance* 3(1): 993860. DOI: 10.1080/23322039.2014.993860.

Halket J, Nesheim L and Oswald F (2015) *The housing stock, housing prices, and user costs: The roles of location, structure and unobserved quality*. London, UK: Cemmap working paper, Centre for Microdata Methods and Practice. DOI: 10.1920/wp.cem.2015.7315.

Hamnett C (1983) Regional Variations in House Prices and House Price Inflation 1969-81. *Area* 15(2): 97–109.

Hamnett C and Reades J (2019) Mind the gap: implications of overseas investment for regional house price divergence in Britain. *Housing Studies* 34(3): 388–406. DOI: 10.1080/02673037.2018.1444151.

Han J (2011) Data Mining: Concepts and Techniques. 3rd edition. Morgan Kaufmann. Available at: https://safarijv.auth0.com/authorize?client_id=UtNi1m1IRXgzYFIwZrhSxell9EDRaL2v&response_type=code&connection=university-college-london&redirect_uri=https://www.safaribooksonline.com/complete/auth0-oauth2/&state=/library/view/-/9780123814791/?ar.

Helbich M, Brunauer W, Vaz E, et al. (2014) Spatial Heterogeneity in Hedonic House Price Models: The Case of Austria. *Urban Studies* 51(2): 390–411. DOI: 10.1177/0042098013492234.

Hiber C (2013) Help to Buy will likely have the effect of pushing up house prices further, making housing become less – not more – affordable for young would-be-owners. London School of Economics and Political Science. Available at: http://blogs.lse.ac.uk/politicsandpolicy (accessed 5 September 2021).

Hilber CAL and Schöni O (2016) Housing Policies in the United Kingdom, Switzerland, and the United States: Lessons Learned. *Cityscape* 18(3). US Department of Housing and Urban Development: 291–332.

HM Land Registry (2015) Additional Price Paid Data release improves market insight. Available at: https://www.gov.uk/government/news/additional-price-paid-data-release-improves-market-insight (accessed 22 November 2018).

HM Land Registry (2016) How to access HM Land Registry Price Paid Data. Available at: https://www.gov.uk/guidance/about-the-price-paid-data#data-excluded-from-price-paid-data (accessed 20 January 2020).

Hodson T (2019) Bracknell Property Price Forecast to Be Fastest Growing Region in 2019. In: *SevenCapital*. Available at: https://sevencapital.com/property-news/bracknell-property-price-forecast-2019/ (accessed 9 January 2020).

Holland R (2019) The Berkshire tech hotspot where house prices are set to soar. Available at: https://www.homesandproperty.co.uk/property-news/buying-in-bracknell-the-berkshire-tech-hotspot-where-house-prices-are-set-to-soar-a129561.html (accessed 9 January 2020).

Holly S, Hashem Pesaran M and Yamagata T (2011) The spatial and temporal diffusion of house prices in the UK. *Journal of Urban Economics* 69(1): 2–23. DOI: 10.1016/j.jue.2010.08.002.

Hox J (2017) *Multilevel Analysis: Techniques and Applications* (eds M Moerbeek and R van de Schoot). Third Edition. New York, NY: Routledge.

Huang B, Wu B and Barry M (2010) Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *International Journal of Geographical Information Science* 24(3): 383–401. DOI: 10.1080/13658810802672469.

Hudson C, Hudson J and Morley B (2018) Differing house price linkages across UK regions: A multi-dimensional recursive ripple model. *Urban Studies* 55(8): 1636–1654. DOI: 10.1177/0042098017700804.

Hudson N (2018) *A housing crisis?* 15 October. Available at: http://resi-analysts.com/wp-content/uploads/2018/10/A-Housing-Crisis-2018-10-15.pdf (accessed 25 November 2018).

Hügel S (2017) Polylabel_cmd: A Command-Line Utility for Generating Optimum Polygon Label Coordinates. Rust. Available at: https://github.com/urschrei/polylabel_cmd (accessed 29 January 2018).

Hulchanski JD (1995) The concept of housing affordability: Six contemporary uses of the housing expenditure-to-income ratio. *Housing Studies* 10(4): 471–491. DOI: 10.1080/02673039508720833.

Iacoviello M and Neri S (2010) Housing Market Spillovers: Evidence from an Estimated DSGE Model. *American Economic Journal: Macroeconomics* 2(2): 125–164. DOI: 10.1257/mac.2.2.125.

Ismail N (2018) The biggest tech hubs in the UK - and which is right for your business? In: *Information Age*. Available at: https://www.information-age.com/biggest-tech-hubs-uk-right-business-123472568/ (accessed 28 February 2020).

Jain AK (2010) Data clustering: 50 years beyond K-means. *Pattern Recognition Letters* 31(8). Award winning papers from the 19th International Conference on Pattern Recognition (ICPR): 651–666. DOI: 10.1016/j.patrec.2009.09.011.

Jeffreys H (1946) An Invariant Form for the Prior Probability in Estimation Problems. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* 186(1007): 453–461.

Jennings M (2018) *Comparing house price indices in the UK*. Available at: https://www.gov.uk/government/publications/about-the-uk-house-price-index/comparing-house-price-indices-in-the-uk (accessed 18 April 2018).

John B (2015) What is ‘affordable housing’? Available at: http://blog.shelter.org.uk/2015/08/what-is-affordable-housing/ (accessed 26 February 2018).

Joice P (2014) Measuring Housing Affordability. *Cityscape; Washington* 16(1): 299–307.

Jones C, Watkins C and Watkins D (2011) Measuring local affordability: variations between housing market areas. *International Journal of Housing Markets and Analysis; Bingley* 4(4): 341–356. DOI: http://dx.doi.org/10.1108/17538271111172148.

Jones K (1991a) *Multi-Level Models for Geographical Research*. Concepts and techniques in modern geography. Norwich: Environmental Publications. Available at: http://www.ggy.bris.ac.uk/personal/KelvynJones/54-multi-level-models.pdf.

Jones K (1991b) Specifying and Estimating Multi-Level Models for Geographical Research. *Transactions of the Institute of British Geographers* 16(2): 148–159. DOI: 10.2307/622610.

Jones K and Bullen N (1993) A Multi-level Analysis of the Variations in Domestic Property Prices: Southern England, 1980-87. *Urban Studies* 30(8): 1409–1426. DOI: 10.1080/00420989320081341.

Jones K and Bullen N (1994) Contextual models of urban house prices: A comparison of fixed- and random-coefficient models developed by expansion. *Economic Geography; Oxford* 70(3): 252–272.

Jones R (2017) Goodbye 25-year mortgages, but are we walking into a borrowing trap? *the Guardian*, 29 July. Available at: http://www.theguardian.com/money/2017/jul/29/goodbye-25-year-mortgage-borrowing-trap-35-year-deals (accessed 1 February 2021).

Kain JF and Quigley JM (1970) Measuring the Value of Housing Quality. *Journal of the American Statistical Association* 65(330): 532–548. DOI: 10.2307/2284565.

Kane TJ, Riegg SK and Staiger DO (2006) School Quality, Neighborhoods, and Housing Prices. *American Law and Economics Review* 8(2): 183–212. DOI: 10.1093/aler/ahl007.

Kassambara MA (2017) *Practical Guide to Cluster Analysis in R: Unsupervised Machine Learning: Volume 1*. 1 edition. Frankreich: CreateSpace Independent Publishing Platform.

Khan SS and Ahmad A (2004) Cluster center initialization algorithm for K-means clustering. *Pattern Recognition Letters* 25(11): 1293–1302. DOI: 10.1016/j.patrec.2004.04.007.

Kiel KA and Zabel JE (2008) Location, location, location: The 3L Approach to house price determination. *Journal of Housing Economics* 17(2): 175–190. DOI: 10.1016/j.jhe.2007.12.002.

Knoll K, Schularick M and Steger T (2017) No Price Like Home: Global House Prices, 1870-2012. *American Economic Review* 107(2): 331–353. DOI: 10.1257/aer.20150501.

Kollewe J (2017) UK rents expected to rise faster than house prices in next five years. Available at: http://www.theguardian.com/business/2017/feb/09/uk-rents-rise-faster-house-prices-next-five-years-rics-survey (accessed 24 March 2018).

Kutty NK (2005) A new measure of housing affordability: Estimates and analytical results. *Housing Policy Debate* 16(1): 113–142. DOI: 10.1080/10511482.2005.9521536.

Lamont O and Stein JC (1999) Leverage and house-price dynamics in U.S. cities. *RAND journal of economics* 30(3): 498–514.

Lancaster KJ (1966) A New Approach to Consumer Theory. *Journal of Political Economy* 74(2): 132.

Laura G and Vidhya A (2014) The Home Stretch: coping with high housing costs. In: *Resolution Foundation*. Available at: http://www.resolutionfoundation.org/app/uploads/2014/12/The-home-stretch1.pdf (accessed 24 March 2018).

Law S (2017) Defining Street-based Local Area and measuring its effect on house price using a hedonic price approach: The case study of Metropolitan London. *Cities* 60: 166–179. DOI: 10.1016/j.cities.2016.08.008.

Law S (2018) *A multi-scale exploration of the relationship between spatial network configuration and housing prices using the hedonic price approach. A Greater London case study*. PhD Thesis. UCL (University College London), London. Available at: http://discovery.ucl.ac.uk/10041030/ (accessed 29 September 2018).

Leishman C (2009) Spatial Change and the Structure of Urban Housing Sub-markets. *Housing Studies* 24(5): 563–585. DOI: 10.1080/02673030903082310.

Leung Y, Mei C-L and Zhang W-X (2000) Statistical Tests for Spatial Nonstationarity Based on the Geographically Weighted Regression Model. *Environment and Planning A: Economy and Space* 32(1): 9–32. DOI: 10.1068/a3162.

Levina I, Sturrock R, Varadi A, et al. (2019) *Modelling the distribution of mortgage* *debt*. *Bank of England working papers*. Bank of England working papers 808, 19 July. Bank of England. Available at: https://ideas.repec.org/p/boe/boeewp/0808.html (accessed 11 March 2021).

Leyland AH and Groenewegen PP (2003) Multilevel modelling and public health policy. *Scandinavian Journal of Public Health* 31(4): 267–274. DOI: 10.1080/14034940210165028.

Li MM and Brown HJ (1980) Micro-Neighborhood Externalities and Hedonic Housing Prices. *Land Economics* 56(2): 125–141. DOI: 10.2307/3145857.

Li Z, Fotheringham AS, Li W, et al. (2019) Fast Geographically Weighted Regression (FastGWR): a scalable algorithm to investigate spatial process heterogeneity in millions of observations. *International Journal of Geographical Information Science* 33(1). Taylor & Francis: 155–175. DOI: 10.1080/13658816.2018.1521523.

Löchl M and Axhausen KW (2010) Modeling hedonic residential rents for land use and transport simulation while considering spatial effects. *Journal of Transport and Land Use* 3(2): 39–63.

Lu B, Charlton M and Fotheringhama AS (2011) Geographically Weighted Regression Using a Non-Euclidean Distance Metric with a Study on London House Price Data. *Procedia Environmental Sciences* 7. Spatial Statistics 2011: Mapping Global Change: 92–97. DOI: 10.1016/j.proenv.2011.07.017.

Lu B, Charlton M, Harris P, et al. (2014) Geographically weighted regression with a non-Euclidean distance metric: a case study using hedonic house price data. *International Journal of Geographical Information Science* 28(4): 660–681. DOI: 10.1080/13658816.2013.865739.

Lund B (2017) *Understanding Housing Policy (Third Edition)*. Bristol: Policy Press.

MacDonald R and Taylor MP (1993) Regional House Prices in Britain: Long-Run Relationships and Short-Run Dynamics. *Scottish Journal of Political Economy* 40(1): 43–55. DOI: 10.1111/j.1467-9485.1993.tb00636.x.

Mace A, Blanc F, Gordon IR, et al. (2016) A 21st century metropolitan green belt. Available at: http://eprints.lse.ac.uk/68012/ (accessed 9 January 2020).

Marsden J (2015) House prices in London – an economic analysis of London's housing market. November. Greater London Authority. Available at: https://www.london.gov.uk/sites/default/files/house-prices-in-london.pdf (accessed 21 November 2016).

Maslow AH (1943) A theory of human motivation. *Psychological Review* 50(4): 370–396. DOI: 10.1037/h0054346.

McLeod S (2018) Maslow's Hierarchy of Needs. In: *Simply Psychology*. Available at:

https://www.simplypsychology.org/maslow.html (accessed 25 November 2018).

Meen G (1999) Regional House Prices and the Ripple Effect: A New Interpretation. *Housing Studies* 14(6): 733–753. DOI: 10.1080/02673039982524.

Meen G (2018) How should affordability be measured? Available at: https://housingevidence.ac.uk/publications/how-should-affordability-be-measured/ (accessed 5 January 2021).

Merlo J, Chaix B, Yang M, et al. (2005) A brief conceptual tutorial of multilevel analysis in social epidemiology: linking the statistical concept of clustering to the idea of contextual phenomenon. *Journal of Epidemiology & Community Health* 59(6): 443–449. DOI: 10.1136/jech.2004.023473.

MHCLG (2015) Technical housing standards – nationally described space standard. Available at: https://www.gov.uk/government/publications/technical-housing-standards-nationally-described-space-standard (accessed 22 April 2021).

MHCLG (2019) Public attitudes to housing: British Social Attitudes survey. 20 July. Available at: https://www.gov.uk/government/collections/public-attitudes-to-housing-british-social-attitudes-survey (accessed 5 September 2021).

MHCLG (2020) 'Permitted development' homes to meet space standards. Available at: https://www.gov.uk/government/news/permitted-development-homes-to-meet-space-standards (accessed 21 April 2021).

Miller N, Peng L and Sklarz M (2011) House Prices and Economic Growth. *The Journal of Real Estate Finance and Economics* 42(4): 522–541. DOI: 10.1007/s11146-009-9197-8.

Millins D and Murie A (2006) *Housing Policy in the UK*. 2006 edition. Basingstoke: Palgrave.

Mok HMK, Chan PPK and Cho Y-S (1995) A hedonic price model for private properties in Hong Kong. *The Journal of Real Estate Finance and Economics* 10(1): 37–48. DOI: 10.1007/BF01099610.

Morancho AB (2003) A hedonic valuation of urban green areas. *Landscape and Urban Planning* 66(1): 35–41. DOI: 10.1016/S0169-2046(03)00093-8.

Morley B and Thomas D (2016) An Empirical Analysis of UK House Price Risk Variation by Property Type. *Review of Economics & Finance* 6: 45–56.

Morphet J and Clifford B (2020) *Reviving Local Authority Housing Delivery: Challenging Austerity Through Municipal Entrepreneurialism*. Bristol: Policy Press.

Mulliner E, Malys N and Maliene V (2016) Comparative analysis of MCDM methods for the assessment of sustainable housing affordability. *Omega* 59: 146–156.

DOI: 10.1016/j.omega.2015.05.013.

Murie A (2016) *The Right to Buy?* 1 edition. Bristol, UK: Policy Press.

Narayan S and Narayan PK (2011) The Importance of Real and Nominal Shocks on the UK Housing Market. *International Journal of Business and Economics; Taichung* 10(3): 219–234.

Nationwide Building Society (2019) House price index. Available at: https://www.nationwide.co.uk/house-price-index#xtab:uk-series. (accessed 4 September 2019).

Nationwide Building Society (2015) Nationwide House Price Index Methodoldy. July. Available at: https://www.nationwide.co.uk/-/media/MainSite/documents/about/house-price-index/nationwide-hpi-methodology.pdf (accessed 4 November 2018).

Nationwide Building Society (2017) Nationwide House Price Index Methodoldy. Available at: https://www.nationwide.co.uk/about/house-price-index/headlines (accessed 4 November 2018).

Nielsen F (2010) A family of statistical symmetric divergences based on Jensen's inequality. Available at: https://arxiv.org/abs/1009.4004 (accessed 8 October 2018).

Office for National Statistics, Land Registry, Registers of Scotland and Land & Property, et al. (2016) *Development of a single Official House Price Index*. 1 February. Available at: https://www.ons.gov.uk/economy/inflationandpriceindices/methodologies/developmentofasingleofficialhousepriceindex (accessed 10 October 2018).

ONS (2010) National Statistician's Review of Housing Market Statistics, 2010. Available at: https://www.statisticsauthority.gov.uk/archive/national-statistician/ns-reports--reviews-and-guidance/national-statistician-s-reviews/national-statistician-s-review-of-housing-market-statistics.html (accessed 19 November 2018).

ONS (2012a) Live tables on housing market and house prices. Available at: https://www.gov.uk/government/statistical-data-sets/live-tables-on-housing-market-and-house-prices (accessed 26 November 2018).

ONS (2012b) National Statistician's Review of Housing Market Statistics, 2012. December. Available at: https://www.statisticsauthority.gov.uk/archive/national-statistician/ns-reports--reviews-and-guidance/national-statistician-s-reviews/national-statistician-s-review-of-housing-market-statistics.html (accessed 4 November 2018).

ONS (2013) Official house price statistics explained. 4 April. DOI: http://webarchive.nationalarchives.gov.uk/20151014022710/http://www.ons.g ov.uk.

ONS (2015a) House Price Statistics for Small Areas in England and Wales: 1995 to 2013. 17 February. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/housing/articles/housepricestatisticsforsmallareasinenglandandwales/2015-02-17 (accessed 18 February 2019).

ONS (2015b) Small Area Model-Based Income Estimates: 2011/12. 20 October. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/bulletins/smallareamodelbasedincomeestimates/2015-10-20 (accessed 17 March 2018).

ONS (2016) House price statistics for small areas in England and Wales: year ending September 2015. 16 June. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/housing/bulletins/housepricestatisticsforsmallareas/yearendingdec1995toyearendingsept2015 (accessed 18 November 2018).

ONS (2017a) Family spending in the UK: financial year ending March 2016. 16 February. Available at: https://www.ons.gov.uk/releases/familyspendingintheuk2016 (accessed 27 May 2018).

ONS (2017b) House price per square metre and house price per room. 11 October. Available at: https://www.ons.gov.uk/economy/inflationandpriceindices/articles/housepricepersquaremetreandhousepriceperroomenglandandwales/2004to2016 (accessed 22 March 2018).

ONS (2017c) House price statistics for small areas in England and Wales: year ending June 2017. December. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/housing/bulletins/housepricestatisticsforsmallareas/yearendingjune2017 (accessed 4 November 2018).

ONS (2017d) House prices: how much does one square metre cost in your area? 10 November. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/housing/articles/housepriceshowmuchdoesonesquaremetrecostinyourarea/2017-10-11 (accessed 30 May 2018).

ONS (2017e) Housing affordability in England and Wales: 1997 to 2016. 17 March. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/housing/bulletins/housingaffordabilityinenglandandwales/1997to2016 (accessed 17 March 2018).

ONS (2017f) *Public value of Statistics on Housing and Planning in the UK*. November.

Available at: https://www.statisticsauthority.gov.uk/publication/public-value-of-statistics-on-housing-and-planning-in-the-uk/ (accessed 4 November 2018).

ONS (2017g) UK House Price Index: December 2016. Available at: https://www.ons.gov.uk/economy/inflationandpriceindices/bulletins/housepriceindex/dec2016 (accessed 5 September 2021).

ONS (2018a) First-time buyer housing affordability in England and Wales. 15 July. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/housing/articles/firsttimebuyerhousingaffordabilityinenglandandwales/2017 (accessed 18 November 2018).

ONS (2018b) House price statistics for small areas Quality and Methodology Information. 24 September. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/housing/methodologies/housepricestatisticsforsmallareasqmi (accessed 4 November 2018).

ONS (2018c) Research Outputs: Rent price statistics for small areas, Great Britain. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/housing/articles/researchoutputsrentpricestatisticsforsmallareasgreatbritain/2012to2016 (accessed 29 April 2018).

ONS (2020a) Housing affordability ratios for Middle layer Super Output areas. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/housing/datasets/housingaffordabilityratiosformiddlelayersuperoutputareas (accessed 5 January 2021).

ONS (2020b) Income estimates for small areas in England and Wales, technical report: financial year ending 2018. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/methodologies/incomeestimatesforsmallareasinenglandandwalestechnicalreportfinancialyearending2018 (accessed 6 January 2021).

ONS (2020c) Research Output: Alternative measures of housing affordability. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/housing/articles/alternativemeasuresofhousingaffordability/financialyearending2018 (accessed 5 January 2021).

Orford S (2002) Valuing Locational Externalities: A GIS and Multilevel Modelling Approach. *Environment and Planning B: Planning and Design* 29(1): 105–127. DOI: 10.1068/b2780.

Orford S (2010) Towards a Data-Rich Infrastructure for Housing-Market Research: Deriving Floor-Area Estimates for Individual Properties from Secondary Data

Sources. *Environment and Planning B: Planning and Design* 37(2): 248–264. DOI: 10.1068/b35082.

Orford S (2017) *Valuing the Built Environment : GIS and House Price Analysis*. London: Routledge. DOI: 10.4324/9781315235134.

Osborne H (2014) UK house price rises for 2014 almost twice as high as predicted. *The Guardian*, 26 August. Available at: https://www.theguardian.com/money/2014/aug/26/uk-house-prices-rises-twice-predicted (accessed 18 November 2018).

Osborne H (2016) Reading takes the lead in house price rises. *The Guardian*, 29 January. Available at: https://www.theguardian.com/money/2016/jan/29/house-prices-rental-costs-continue-rise-ons-land-registry (accessed 11 January 2020).

Osborne H and Neate R (2021) UK hits record number of homes sold in a month. Available at: http://www.theguardian.com/business/2021/jul/21/uk-record-houses-sold-in-a-month-stamp-duty (accessed 12 September 2021).

Osland L and Thorsen I (2008) Effects on Housing Prices of Urban Attraction and Labor-Market Accessibility. *Environment and Planning A: Economy and Space* 40(10): 2490–2509. DOI: 10.1068/a39305.

Palm R (1978) Spatial Segmentation of the Urban Housing Market. *Economic Geography* 54(3): 210–221. DOI: 10.2307/142835.

Partington R and Perraudin F (2018) London slump shrinks north-south divide in house prices. *The Guardian*, 16 March. Available at: https://www.theguardian.com/business/2018/mar/16/london-slump-shrinks-north-south-divide-in-house-prices (accessed 1 November 2018).

Peachey K (2017) House price ‘north-south divide’ narrows. 14 November. Available at: https://www.bbc.com/news/business-41981342 (accessed 1 November 2018).

Powell-Smith A (2017) House prices by square metre in England & Wales. Available at: https://houseprices.anna.ps (accessed 19 November 2018).

Pryce P, Corry D and Beatson B (2011) Investment in housing and its contribution to economic growth - Shelter England. 4 October. Available at: http://england.shelter.org.uk/professional_resources/policy_and_research/policy_library/policy_library_folder/investment_in_housing_and_its_contribution_to_economic_growth (accessed 20 August 2018).

Rae A (2015) Online Housing Search and the Geography of Submarkets. *Housing Studies* 30(3): 453–472. DOI: 10.1080/02673037.2014.974142.

Rae A (2017) The Geography of Travel to Work in England and Wales: Extracts from the 2011 Census. *Applied Spatial Analysis and Policy* 10(4): 457–473. DOI: 10.1007/s12061-016-9196-0.

Raudenbush SW and Bryk AS (2002) *Hierarchical Linear Models: Applications and Data Analysis Methods*. 2nd Edition. Advanced quantitative techniques in the social sciences ; 1. Thousand Oaks, CA: Sage Publications.

Reinold K (2011) *Housing Equity Withdrawal Since the Financial Crisis*. ID 1865200, SSRN Scholarly Paper, 13 June. Rochester, NY: Social Science Research Network. Available at: https://papers.ssrn.com/abstract=1865200 (accessed 20 August 2018).

Ren Z, Wang J, Liao Y, et al. (2013) Using spatial multilevel regression analysis to assess soil type contextual effects on neural tube defects. *Stochastic Environmental Research and Risk Assessment* 27(7): 1695–1708. DOI: 10.1007/s00477-013-0707-0.

Richardson HW (2013) *The New Urban Economics: And Alternatives*. 1st edition. Oxon: Routledge.

RICS (2018) UK Residential Market Survey. Available at: https://www.rics.org/uk/knowledge/market-analysis/rics-residential-market-survey/ (accessed 27 February 2018).

Robbie de Santos (2011) Briefing: Generation Rent - learning from different rental markets. November. Available at: http://england.shelter.org.uk/professional_resources/policy_and_research/policy_library/policy_library_folder/briefing_generation_rent_-_learning_from_different_rental_markets (accessed 24 March 2018).

Rohde N (2016) J-divergence measurements of economic inequality. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 179(3): 847–870. DOI: 10.1111/rssa.12153.

Rosen S (1974) Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy* 82(1): 34.

Scanlon K, Whitehead CME and Blanc F (2017) *A taxing question: is Stamp Duty Land Tax suffocating the English housing market?* Monograph, November. Available at: http://lselondonhousing.org/ (accessed 11 November 2019).

Sedgwick P (2011) The ecological fallacy. *BMJ* 343: d4670. DOI: 10.1136/bmj.d4670.

Shelter (2015) Housing affordability for first time buyers. Available at: https://england.shelter.org.uk/professional_resources/policy_and_research/policy_library/policy_library_folder/report_affordability_for_first_time_buyers (accessed 19 September 2018).

Shen Y and Karimi K (2015) Understanding the roles of urban configuration on spatial heterogeneity and submarket regionalisation of house price pattern in a mix-scale hedonic model: The case of Shanghai, China. In: *The 10th Space Syntax Symposium (SSS10)*, london, 2015, pp. 1–19. UCL. Available at: https://discovery.ucl.ac.uk/id/eprint/1498769/1/SSS10_Proceedings_098.pdf.

Singer JD and Willett JB (2003) *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. Oxford scholarship online. New York ; Oxford: Oxford University Press. Available at: http://dx.doi.org/10.1093/acprof:oso/9780195152968.001.0001 (accessed 4 January 2020).

Sirmans GS, MacDonald L, Macpherson DA, et al. (2006) The Value of Housing Characteristics: A Meta Analysis. *The Journal of Real Estate Finance and Economics* 33(3): 215–240. DOI: 10.1007/s11146-006-9983-5.

Sirmans S, Macpherson D and Zietz E (2005) The Composition of Hedonic Pricing Models. *Journal of Real Estate Literature* 13(1): 1–44. DOI: 10.5555/reli.13.1.j03673877172w0w2.

Smith N (1987) Gentrification and the Rent Gap. *Annals of the Association of American Geographers* 77(3): 462–465. DOI: 10.1111/j.1467-8306.1987.tb00171.x.

Smith NB (2018) We used big data to predict house prices – and this is what we found. Available at: https://www.architectsjournal.co.uk/opinion/we-used-big-data-to-predict-house-prices-and-this-is-what-we-found/10024129.article (accessed 1 November 2018).

Smith S (2017) The university towns where house prices have risen an average 22pc in three years. *The Telegraph*, 23 September. Available at: https://www.telegraph.co.uk/property/house-prices/university-towns-house-prices-have-risen-average-22pc-three/ (accessed 25 January 2020).

South B and Henretty N (2017) House price statistics for small areas: Using administrative data to give new insights. *Statistical Journal of the IAOS* 33(3): 609–614. DOI: 10.3233/SJI-160340.

Steele F (2008a) Module 5: Introduction to Multilevel Modelling (Concepts). Centre for Multilevel Modelling, University of Bristol. Available at: https://research-information.bris.ac.uk/en/publications/module-5-introduction-to-multilevel-modelling-concepts (accessed 8 December 2020).

Steele F (2008b) Multilevel models for longitudinal data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 171(1): 5–19. DOI: 10.1111/j.1467-985X.2007.00509.x.

Stephens M (2012) Tackling Housing Market Volatility in the UK. Part I: Long- and Short-term Volatility. *International Journal of Housing Policy* 12(3): 367–380. DOI: 10.1080/14616718.2012.709674.

Stevenson S (2004) House price diffusion and inter-regional and cross-border house price dynamics. *Journal of Property Research* 21(4): 301–320. DOI: 10.1080/09599910500151228.

Stone ME (2006a) A Housing Affordability Standard for the UK. *Housing Studies* 21(4): 453–476. DOI: 10.1080/02673030600708886.

Stone ME (2006b) What is housing affordability? The case for the residual income approach. *Housing Policy Debate* 17(1): 151–184. DOI: 10.1080/10511482.2006.9521564.

Stone ME, Burke T, Ralston L, et al. (2011) *The Residual Income Approach to Housing Affordability: The Theory and the Practice*. Melbourne: Australian Housing and Urban Research Institute.

Swank J, Kakes J and Tieman AF (2003) The housing ladder, taxation, and borrowing constraints. Available at: https://repub.eur.nl/pub/808/ (accessed 15 July 2018).

Szumilo N, Laszkiewicz E and Fuerst F (2017) The spatial impact of employment centres on housing markets. *Spatial Economic Analysis* 12(4): 472–491. DOI: 10.1080/17421772.2017.1339119.

Thwaites G and Wood R (2005) *The Measurement of House Prices*. ID 707043, SSRN Scholarly Paper, 27 April. Rochester, NY: Social Science Research Network. Available at: https://papers.ssrn.com/abstract=707043 (accessed 17 October 2018).

Tsenkova S and French M (2011) *Affordable Land and Housing in Europe and North America*. UN-HABITAT. Available at: https://unhabitat.org/books/affordable-land-and-housing-in-europe-and-north-america-2/ (accessed 14 February 2018).

UN-HABITAT (2011) Affordable Land and Housing in Asia. Available at: https://unhabitat.org/books/affordable-land-and-housing-in-asia-2/ (accessed 12 April 2018).

UN-HABITAT (2012) Affordable Land and Housing in Africa. Available at: https://unhabitat.org/books/affordable-land-and-housing-in-africa/ (accessed 12 April 2018).

Vincent M (2020) House prices could fall by up to 5% in 2021. Available at: https://www.dailymail.co.uk/news/article-9078609/House-prices-fall-5-2021-wipe-12-660-value-average.html (accessed 12 September 2021).

Visser P, Dam FV and Hooimeijer P (2008) Residential Environment and Spatial Variation in House Prices in the Netherlands. *Tijdschrift voor economische en sociale geografie* 99(3): 348–360. DOI: 10.1111/j.1467-9663.2008.00472.x.

Wang Y, Potoglou D, Orford S, et al. (2015) Bus stop, property price and land value tax: A multilevel hedonic analysis with quantile calibration. *Land Use Policy* 42: 381–391. DOI: 10.1016/j.landusepol.2014.07.017.

Whitehead C and Williams P (2011) Causes and Consequences? Exploring the Shape and Direction of the Housing System in the UK Post the Financial Crisis. *Housing Studies* 26(7–8). Routledge: 1157–1169. DOI: 10.1080/02673037.2011.618974.

Whitehead C, Monk S, Clarke A, et al. (2008) Measuring Housing Affordability: A Review of Data Sources.: 98.

Whitehead CME (1991) From Need to Affordability: An Analysis of UK Housing Objectives. *Urban Studies* 28(6): 871–887. DOI: 10.1080/00420989120081101.

Wood R (2015) A comparison of UK residential house price indices. (21): 212–227.

Wu B, Li R and Huang B (2014) A geographically and temporally weighted autoregressive model with application to housing prices. *International Journal of Geographical Information Science* 28(5): 1186–1204. DOI: 10.1080/13658816.2013.878463.

Yao J and Fotheringham AS (2016) Local Spatiotemporal Modeling of House Prices: A Mixed Model Approach. *The Professional Geographer* 68(2): 189–201. DOI: 10.1080/00330124.2015.1033671.

Yu D (2007) Modeling Owner-Occupied Single-Family House Values in the City of Milwaukee: A Geographically Weighted Regression Approach. *GIScience & Remote Sensing* 44(3): 267–282. DOI: 10.2747/1548-1603.44.3.267.

Zaninotto P, Falaschetti E and Sacker A (2009) Age trajectories of quality of life among older adults: results from the English Longitudinal Study of Ageing. *Quality of Life Research* 18(10): 1301–1309. DOI: 10.1007/s11136-009-9543-6.

Appendices

Appendix A1

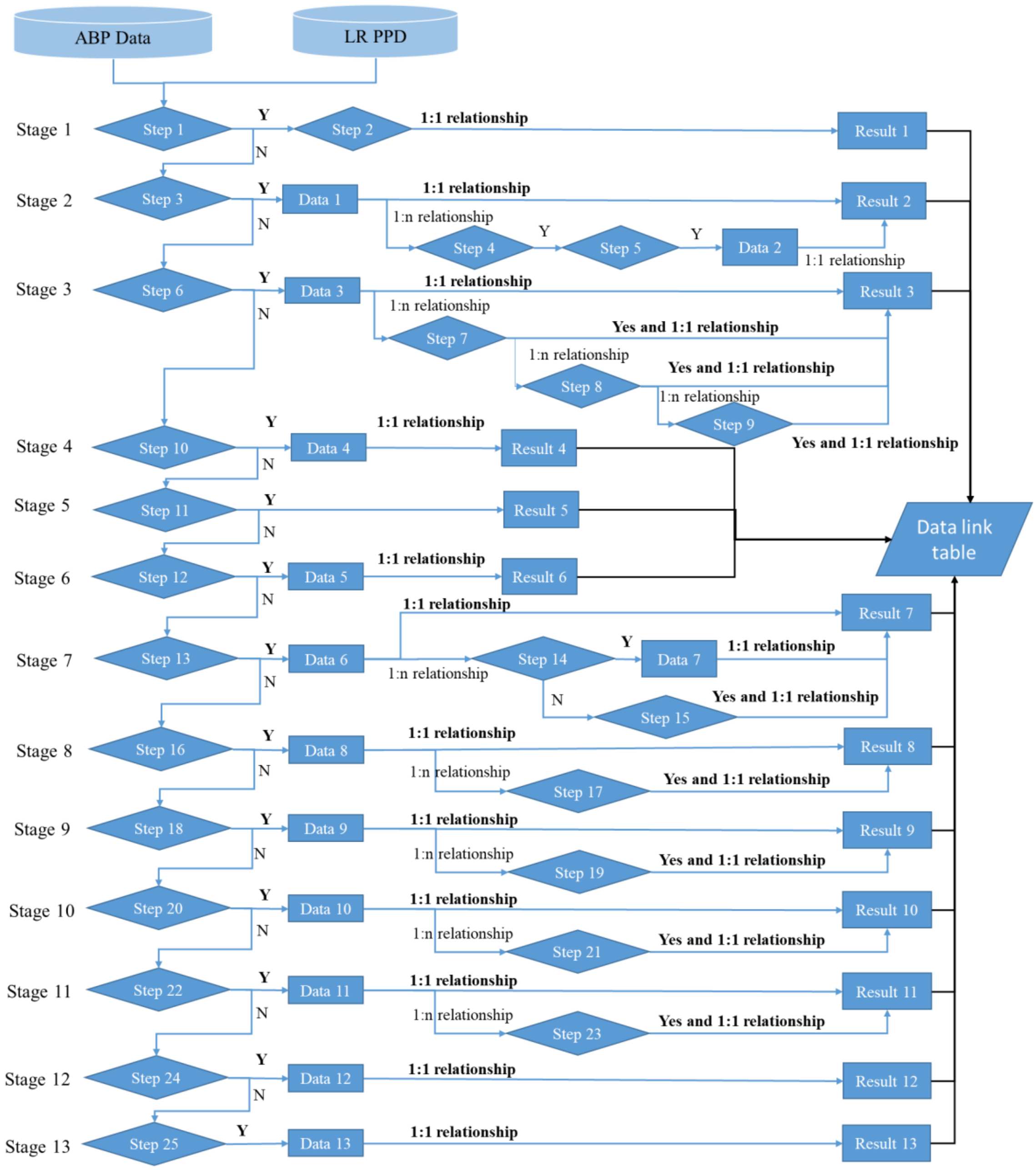


Figure A1 Master workflow of the 13 stages data linkage

Figure A1 presents the master workflow of 13 stages of data linkage between LR PPD and ABP data. Each stage contains more than one step and each step contains more than one match rules. Details of the match rules for each step are listed in Table A1. For each step, we assess whether all corresponding matching rules listed in the table are satisfied. If yes, the matching process will go the branch marked "Y" in Figure A1; otherwise, the matching process follow the branch marked "N". Take the first match part in Stage 2 as an example, in the OS MMTL, TOID[35] is a unique reference to identify the building feature. TOID contained in the ABP data is named as ostopotoid. Meanwhile in the LR PPD, each transaction has a unique identifier named as transactionid. In each step we loop the matching rules within the same postcode. When putting in the Data 1 in the matching process of step 3. Firstly, a function starts with creating a dataset which contains all the unique postcodes from LR PPD (temp data1), then the function continue subset all the records from LR PPD and ABP from a given postcode unit in temp data1, then the match executes the match rules in step 3 (i.e. “test whether PAON of each transaction in LR PPD is equal to buildingname in ABP” or “test whether PAON of each transaction in LR PPD is equal to buildingnumber in ABP” or “test whether PAON of each transaction in LR PPD is equal to bb in ABP”), if the result is YES then transactionid and ostopotoid will directly link based on the match rules in step 3 and restore in Data 1. After this, a new function will be used to identify if there is a one transactionid match one ostopotoid and if the result is YES and this tested link result will store in Result 2 dataset. Otherwise it will go to Stage 3 to conduct the match test in step 4. Following this all the successful 1:1 match link in Stage 2 will store in Result 2 dataset and final store in Data link table. All the matching

[35] All the words coloured in grey shading are the fields name.

process in Figure A1 works the same as described above and the final result is data link table. The data linkage job is conducted in RStudio.

Table A1 Details of 97 matching rules in 13 stages[36]

Stage No.	Step No.	Match rules
Stage 1	Step 1	PAON is NULL and SAON is not NULL
	Step 2	SAON is equal to pp
		SAON is equal to paostartnumber1
		SAON is equal to saostartnumber
		SAON is equal to buildingname
		SAON is equal to paotext
		SAON2 is equal pp
		SAONSTREET is equal to buildingname
Stage 2	Step 3	PAON is equal to buildingname
		PAON is equal to buildingnumber
		PAON is equal to bb
	Step 4	SAON is not NULL
	Step 5	PAON is equal to pp and SAON is equal to saotext
		PAON is equal to buildingname and SAON is equal to saotext
		PAON is equal to buildingnumber and SAON is equal to saotext
		PAON is equal to bb and SAON is equal to saotext
		PAON is equal to buildingname and SAON is equal to subbuildingname
		PAON is equal to buildingnumber and SAON is equal to subbuildingname
		PAON is equal to bb and SAON is equal to subbuildingname
		PAON is equal to pp1 and SAON is equal to saotext
		PAON is equal to paotext and SAON is equal to ss
		PAON is equal to bb and SAON is equal to ss
		PAON is equal to buildingname and FLATSAON is equal to subbuildingname

[36] In all the matching rule of this table, capital word coloured in grey stands for the address field in Land Registry, the capitalized word coloured in grey stands for the address field in ABP data.

Stage No.	Step No.	Match rules
		PAON is equal to paotext and FLATSAON to saotex
		PAON is equal to buildingname and FLATSAON is equal to subbuildingnamenew
		PAON is equal to buildingname and SAON is equal to fss
		PAON is equal to paotext and SAON is equal to fss
		PAON is equal to bb and SAON is equal to fss
		SAONPAON is equal to buildingname
		PAON is equal to pp1 and SAON is equal to flatsao
		PAON is equal to pp1 and SAON is equal to saotext2
		PAON is equal to paotext and SAON is equal to saotext
		PAON is equal to buildingname and SAON is equal to ss
		PAON is equal to bb and FLATSAON is equal to saotext
		PAON is equal to pp1 and SAON is equal to ss
		PAON is equal to subbuildingname and SAONPAON is equal to buildingname
		PAON is equal to pp
		PAON is equal to paotext and SAON is equal to pp
	Step 6	SAON is NULL and PAON is equal to buildingnumber
		SAON is NULL and PAON is equal to paotext
Stage 3	Step 7	PAON is equal to paostartnumber
	Step 8	PAON is equal to paostartnumber and SAON is equal to flatpao
	Step 9	PAON is equal to pp and SAON is equal to saotext
		PAON is equal to pp and FLATSAON is equal to saotext
		PAON is equal to pp
		PAON is equal to pp and STREET is equal to streetdescription
Stage 4	Step 10	PAON is equal to pp
Stage 5	Step 11	Direct match when there is only one ostopotoid in its postcode unit
Stage 6	Step 12	PAON is equal to psao
Stage 7	Step 13	PAON is equal to pp2
		PAON is equal to pp1
	Step 14	PAON is equal to pp2 and SAON is equal to saotext
		PAON is equal to pp2 and SAON is equal to ss
		PAON is equal to pp2 and SAON is equal to flatss

Stage No.	Step No.	Match rules
		PAON is equal to pp2 and FLATSAON is equal to saotext
		PAON is equal to pp2 and SAON is equal to unitss
		PAON is equal to pp2 and SAON is equal to subbuildingname
		PAON1 is equal to buildingname and SAON is equal to subbuildingname
	Step 15	Detached, semi-detached and terrace transactions: PAON is equal to pp1 and SAON is equal to ss
		Flat transactions: PAON is equal to pp2 and SAON is equal to ss1
Stage 8	Step 16	PAON is equal to paotext
		PAON is equal to sp
	Step 17	PAON is equal to paotext and SAON is equal to ss
		PAON is equal to paotext and FLATSAON is equal to saotext
		PAON is equal to paotext and SAON is equal to flatss
		PAON is equal to paotext and SAON is equal to saotext
		PAON is equal to paotext and SAON is equal to pp
		PAON is equal to paotext and SAON is equal to subss
		PAON is equal to paotext and SAONPAON is equal to saobui
Stage 9	Step 18	PAON1 is equal to buildingname
		PAON1 is equal to pp4
	Step 19	PAON1 is equal to buildingname and SAON is equal to subbuildingname
		PAON1 is equal to buildingname and SAON is equal to saotext
		PAON2 is equal to pp4 and SAON is NULL
		PAON1 is equal to ppp and SAON is equal to ss
		PAON1 is equal to ppp and SAON is equal to flatss
		PAON1 is equal to ppp and SAON is equal to saotext
		PAON1 is equal to buildingname and FLATSAON is equal to subbuildingname
		PAON1 is equal to buildingname and SAON is equal to flatsub
		PAON1 is equal to buildingname and SAON is equal to ss
		PAON2 is equal to pp4 and SAON is equal to saotext
		PAON2 is equal to pp4 and FLATSAON is equal to saotext
		PAON2 is equal to pp4 and SAON is equal to subbuildingname

Stage No.	Step No.	Match rules
		PAON2 is equal to pp4 and SAON is equal to ssp
Stage 10	Step 20	STREET is equal to paotext
	Step 21	STREET is equal to paotext and PAON is equal to ss
		STREET is equal to paotext and PAON is equal to ss and SAON is equal to saotext and SAON is not NULL
		PAONSTREET is equal to buildingname
		PAONSTREET is equal to paotext
		PAONSTREET is equal to paotext and SAON is equal to ss
		STREET is equal to paotext and FLATPAON is equal to subbuildingname
		STREET is equal to paotext and UNITPAON is equal to saotext
Stage 11	Step 22	PAON is equal to saopp
	Step 23	PAON is equal to saopp and SAON is equal to flatss
Stage 12	Step 24	SAONPAON is equal to buildingname
Stage 13	Step 25	PAON is equal to ss and SAON is NULL
		PAONSTREET is equal to buildingname

Appendix A2

Table A2 New address variables created from LR PPD and EPC datasets

Variable	Create method	Dataset
ADD1	Capitalised the all the string in ADDRESS1, then remove leading and trailing whitespace	Domestic EPCs
ADD2	Capitalised the all the string in ADDRESS2, then remove leading and trailing whitespace	Domestic EPCs
ADD3	Capitalised the all the string in ADDRESS3, then remove leading and trailing whitespace	Domestic EPCs
ADD	Capitalised the all the string in ADDRESS, then remove leading and trailing whitespace	Domestic EPCs
ADD2NEW	Delete the ‘-’ in the ADD2	Domestic EPCs
ADDC	Delete all the ‘/’ , ‘ .’, ‘” punctuation characters and blank space in ADD	Domestic EPCs
ADDU	Delete the ‘UNIT ’ string in the ADD, then delete all the comma and blank space	Domestic EPCs
ADDC3	Delete the comma in ADDC	Domestic

Variable	Create method	Dataset
		EPCs
ADDCC	Delete all the '-' , '/' , ' .', '" punctuation characters and blank space in ADD	Domestic EPCs
ADDCCC	Delete the comma in ADDCC	Domestic EPCs
ADDC4	Delete all the '/' , ' .', '-' punctuation characters and blank space in ADD	Domestic EPCs
ADDC6	Delete all the '" , ' ,' punctuation characters and blank space in ADD	Domestic EPCs
ADDRE	Delete the blank space in ADD	Domestic EPCs
ADDREC	Delete the comma in ADDRE	Domestic EPCs
ADD1C	Delete all the '/' , ' .', '" punctuation characters and blank space in ADD1	Domestic EPCs
ADD1CC	Delete '-' punctuation characters in ADD1C	Domestic EPCs
ADD1C2	Delete the comma in ADD1C	Domestic EPCs
ADD1C3	Delete all the comma and blank space in ADD1	Domestic EPCs
ADD1C6	Delete the 'UNIT ' in ADD1, then delete all the comma and blank space	Domestic EPCs
ADD1C4	Delete '" punctuation characters in ADD1C3	Domestic EPCs
ADD1C5	Delete the '.' and blank space in ADD1	Domestic EPCs
ADD1C7	Delete all the comma and blank space in ADD1	Domestic EPCs
ADD1C8	Delete all the comma in ADD1C5	Domestic EPCs
ADD1C9	Delete the all the blank space in ADD1	Domestic EPCs
ADD1C10	Delete the '/' punctuation characters in ADD1	Domestic EPCs
ADD12C2	Delete the comma in ADD12	Domestic EPCs
ADD12C	Delete ' .', '" ,'/' punctuation characters in ADD12	Domestic EPCs
ADD12C1	Delete ' .', '" ,'/' punctuation characters and comma	Domestic

Variable	Create method	Dataset
	in ADD12	EPCs
ADD12C3	Delete all ‘ .’, ‘” ,‘/’ , ‘-’ punctuation characters and comma in the ADD12	Domestic EPCs
ADD12C4	Delete all the ‘ .’, ‘-’ ,‘/’ and blank space in ADD12	Domestic EPCs
ADD12C5	Delete all the ‘ .’, ‘,’ and blank space in ADD12	Domestic EPCs
ADD13C	Delete ‘ .’, ‘” ,‘/’ punctuation characters and blank space in ADD13	Domestic EPCs
ADD13C1	Delete the comma in ADD13C	Domestic EPCs
ADD13C2	Delete the comma in ADD13	Domestic EPCs
ADD23C	Delete ‘ .’, ‘” ,‘/’ punctuation characters in ADD23	Domestic EPCs
ADD23C1	Delete the comma in ADD23C	Domestic EPCs
ADD161	For the ADD1 contain a comma, then select the text before the first comma	Domestic EPCs
ADD162	For the ADD1 contain a comma and ‘-’ punctuation character, then select the string after the first comma, then delete the ‘-’ punctuation character	Domestic EPCs
ADD165	For the ADD1 contain a comma and ‘.’ punctuation character, then select the string after the first comma, then delete the ‘.’ punctuation character	Domestic EPCs
add1sp	For the ADD2 is not start with number string and also does not contain a word with one character, select the string before the first blank space in ADD1	Domestic EPCs
add63	Delete ‘-’ and ‘.’ in add162	Domestic EPCs
add1nnn	Delete ‘NO ’ string in ADD1, then delete all the comma	Domestic EPCs
ADD1df1	Delete ‘FLAT ’ string in ADD1 , then select the string the first string before the first word boundary, then delete the comma	Domestic EPCs
ADD1du	Delete the ‘UNIT ’ string in ADD1 , then delete all the comma and blank space	Domestic EPCs
ADD163	Select the string before the first blank space in ADD1	Domestic EPCs
add261	For the add2 contain a comma, then select the string before the first comma	Domestic EPCs

Variable	Create method	Dataset
add263	Select the string before the first blank space in ADD2, then delete comma	Domestic EPCs
add31	Delete ‘”, ‘.’ and ‘/’ in ADD3	Domestic EPCs
fladd1c	Delete all the blank space in fladd1	Domestic EPCs
fladdc	Delete all the comma in the fladd	Domestic EPCs
ADD1dff	For the ADD1 has ‘FLAT ‘, delete ‘FLAT ’ string in ADD1	Domestic EPCs
add264	Select the string after the first blank space in ADD2	Domestic EPCs
add2641	Select the string after the first comma in ADD2	Domestic EPCs
apADD1	Delete ‘-’ , ‘/’ , ‘ .’, ‘” ‘,’ punctuation characters and blank space in apadd1	Domestic EPCs
ADDr61	For the ADD contain a comma, then select the string before the first comma	Domestic EPCs
ADDr62	For the ADD contain a comma and -punctuation character, then select the string after the first comma, then delete the ‘-’ , ‘”, ‘.’ and ‘/’ punctuation character	Domestic EPCs
add361	For the ADD3 contain a comma, then select the text before the first comma	Domestic EPCs
ADDC5	Delete all the ‘/’ , ‘ .’ punctuation characters and blank space in ADD	Domestic EPCs
ADDC7	Delete all the ‘-’ punctuation characters and blank space in ADD	Domestic EPCs
ADDC8	Delete all the ‘.’, ‘” punctuation characters and blank space in ADD	Domestic EPCs
ADDC9	Delete all the ‘.’, ‘” and ’/’ punctuation characters in ADD	Domestic EPCs
ADDC10	Delete all the ‘-’ , ‘/’ , ‘ .’, ‘” ,‘,’ punctuation characters and blank space in ADD	Domestic EPCs
ADD262	For the ADD2 contain a comma character, then select the string after the first comma	Domestic EPCs
add1f61	For the ADD1 in EPC data has ‘FLAT ’ string, then delete the FLAT ’ string, then subset the string before the first comma , then delete the all the comma	Domestic EPCs
add1f61f2	Combine ‘FLAT ’ and add1f61 , then combine ADD2 with a comma and a blank space, then delete all the	Domestic EPCs

Variable	Create method	Dataset
	blank space and comma.	
adddap	Delete ‘APARTMENT ’ string in ADD , then delete all the blank space	Domestic EPCs
saonn	Delete all the ‘/’ punctuation characters in SAON	House price spatial data
paonn	Delete all the ‘”, ‘.’ punctuation characters in PAON	House price spatial data
paonn61	Select the string before the first comma in paonn	House price spatial data
paonn2	Delete comma and blank space in PAON	House price spatial data
paonn3	Delete ‘-’ and blank space in PAON	House price spatial data
streetn	Delete all the ‘” punctuation characters in street	House price spatial data
streetn1	Delete ‘-’ , ‘ .’, ‘” punctuation characters and blank space in street	House price spatial data
streetn2	Delete ‘-’ , ‘” punctuation characters and blank space in street	House price spatial data
streetn5	Delete ‘/’ , ‘ .’, ‘” punctuation characters and blank space in street	House price spatial data
localityn	Delete all the ‘”, ‘.’ punctuation characters in locality	House price spatial data
saonpaonstreet31	Delete the comma in saonpaonstreet3	House price spatial data
saonpaonstreetn31	Delete the comma in saonpaonstreetn3	House price spatial data
paon61	For the PAON contain comma, subset the text before the first comma	House price spatial data
paon61c	Delete all the blank space in paon61	House price spatial data
paon62	For the PAON contain comma, subset the string after the first comma	House price spatial data
paon64	Subset the string before the first blank space in PAON	House price spatial data
paon641	Subset the string after the first blank space in PAON	House price spatial data
paon65	For the PAON contain comma, Extract the last word from PAON	House price spatial data

Variable	Create method	Dataset
paon65n	For the paonn contain comma, subset the string after last blank space in paonn	House price spatial data
saon2	Delete ‘APARTMENT ’ string in SAON	House price spatial data
fldsaon	For SAON contain ‘FLAT ’ string and PAON not start with number string. Delete ‘FLAT ’ string in SAON	House price spatial data
fldsaon1	For SAON contain ‘FLAT’ string and PAON start with number string. Delete ‘FLAT ’ string in SAON	House price spatial data
saon7	Replace ‘FLAT ’ string to ‘APARTMENT ’ string in SAON	House price spatial data
saon71	Replace ‘FLAT ’ string to ‘APARTMENT ’ string in saonn	House price spatial data
saonn4	Delete ‘FLAT ’ string in saonn	House price spatial data
saon1	Replace ‘APARTMENT ’ string to ‘FLAT ’ string in saonn	House price spatial data
saonn2	Delete ‘APARTMENT ’ string in saonn	House price spatial data
saonn3	Delete ‘.’ And ‘ / ’ in SAON	House price spatial data
ADD1num	Extract the number string in ADD1	House price spatial data
saonn5	For SAON contain ‘APARTMENT ’ , replace ‘APARTMENT ’ string to ‘UNIT ’ string in SAON and then delete ‘/’ punctuation characters	House price spatial data
sao1	Replace ‘APARTMENT ’ string to ‘FLAT ’ string in SAON	House price spatial data
saon8	Replace ‘LOFT ’ to ‘FLAT ’ in SAON	House price spatial data
saon4	Delete ‘FLAT ’ string in SAON	House price spatial data
paon6164	Select the number string from paon61	House price spatial data
paon6163	Select all the non-digitals from paon61	House price spatial data
paon11	Delete all the comma in the PAON	House price spatial data
ADD12	Combine ADD1 and ADD2 with a comma and a blank space, then delete all the blank space	Domestic EPCs

Variable	Create method	Dataset
ADD12new	Combine ADD1 and add2new with a blank space, then delete the '/' , ' .', '" punctuation characters, blank space and comma	Domestic EPCs
ADD13	Combine ADD1 and ADD3 with a comma and a blank space, then delete all the blank space	Domestic EPCs
ADD23	Combine ADD2 and ADD3 with a blank space, then delete all the blank space	Domestic EPCs
ADD66	Combine ADD161 and ADD162 with a comma and a blank space, then delete all the comma and blank space	Domestic EPCs
ADD662	Combine ADD66 and ADD2 with a comma and a blank space, then delete the comma and blank space	Domestic EPCs
ADD67	Combine ADD161 and ADD165 with a comma and a blank space, then delete all the comma and blank space	Domestic EPCs
ADDSP12	Combine add1sp and add2 with a comma and a blank space, then delete the comma and blank space	Domestic EPCs
ADD68	Combine add161 and add63 with a comma and a blank space, then delete '" and blank space	Domestic EPCs
ADD69	Combine add1nn and ADD2 with a comma and a blank space, then delete all the blank space	Domestic EPCs
ADD1632	Combine ADD163 and ADD2 with a blank space, then delete all the comma and blank space	Domestic EPCs
flADD	Combine 'FLAT ' string and ADD with a comma and a blank space , then delete all the comma and blank space	Domestic EPCs
ADD2611	Combine add261 and add1 with a comma and a blank space, then delete all the comma and blank space	Domestic EPCs
fladd1	Combine 'FLAT ' and ADD1 with a blank space	Domestic EPCs
fladd	Combine 'FLAT ' and ADD with a blank space , then delete all the blank space	Domestic EPCs
flADD13	Combine fladd1 and add31 with a blank space, then delete all the comma and blank space.	Domestic EPCs
ADD5	Combine add263 and ADD1dff, then combine add264, then delete all the blank space	Domestic EPCs
apadd1	Combine 'APARTMENT ' and ADD1 with a blank space	Domestic EPCs
ADDr66	Combine ADDr61 and ADDr62 with a comma and a blank space, then delete all the comma and blank space	Domestic EPCs

Variable	Create method	Dataset
ADD6	Combine ADD1 and ADD2 with a comma and a blank space, then combine add361 with a comma and a blank space, then delete all the, '/' , ' .', '" punctuation characters and blank space	Domestic EPCs
add12643	Combine ADD1 and add2641 with a comma and a blank space, then combine add3 with a comma and a blank space, then delete all the blank spaces	Domestic EPCs
ADD1264	Combine ADD1 and add264 with a comma and a blank space, then delete all the blank space and comma	Domestic EPCs
ADD8	Combine ADD1C10 and ADD2 with a comma and a blank space, then delete all the blank space	Domestic EPCs
ADD7	Combine ADD161 and ADD2 with a blank space, then delete all the blank space	Domestic EPCs
ADD1num2	Combine ADD1num and ADD2 with a comma and a blank space, then delete, '/' , ' .', '" punctuation characters and all blank space	Domestic EPCs
ADD1262	Combine ADD1 and ADD262 with a comma and a blank space, then delete all the blank space	Domestic EPCs
ADD1262C	Combine ADD1 and ADD262 with a comma and a blank space, then delete all the blank space and comma	Domestic EPCs
ADD1262cc	Combine ADD1 and ADD262 with a comma and a blank space, then delete all the blank space and '"	Domestic EPCs
apadd1632	Combine 'APARTMENT' and add163 with a blank space, then combine with ADD2 with a comma and a blank space, then delete all the blank space and comma	Domestic EPCs
saonpaonstreet	Combine SAON and PAON with a comma and a blank space, then combine street with a blank space, then delete all the blank space	House price spatial data
saonpaonstreet5	Combine SAON and PAON with a comma and a blank space, then combine street with a blank space, then delete all the blank space and comma	House price spatial data
saonpaonstreet1	Combine SAON and PAON with a comma and a blank space, then combine street with a comma and a blank space, then delete all the blank space	House price spatial data
saonpaonstreet2	Combine SAON and PAON with a blank space and then remove leading and trailing whitespace, then combine street with a comma and a blank space, then delete all the blank space	House price spatial data
saonpaonstreetn	Combine saonn and paonn with a comma and a blank space, then combine streetn with a blank space, then	House price spatial data

Variable	Create method	Dataset
	delete all the blank space	
saonpaonstreetn1	Combine saonn and paonn with a comma and a blank space, then combine streetn with a comma and a blank space, then delete all the blank space	House price spatial data
saonpaonstreetn2	Combine saonn and paonn with a blank space, then combine streetn with a comma and a blank space, then delete all the blank space	House price spatial data
saonpaonlo	Combine SAON and PAON with a blank space, then combine locality with a comma and a blank space, then delete all the blank space	House price spatial data
saonpaonlon	Combine saonn and paonn with a blank space, then combine localityn with a comma and a blank space, then delete all the blank space	House price spatial data
saonpaonstreet3	Combine SAON and PAON with a blank space, then delete combine street with a blank space, then delete all the blank space	House price spatial data
saonpaonstreetn3	Combine saonn and paonn with a blank space, then delete combine streetn with a blank space, then delete all the blank space	House price spatial data
saonpaonstreetlo	Combine SAON and PAON with a comma and a blank space, then combine street with a comma and a blank space, then combine locality with a comma and a blank space, then delete all the blank space	House price spatial data
saonpaonstreetnlo	Combine saonn and paonn with a comma and a blank space, then combine streetn with a comma and a blank space, then combine localityn with a comma and a blank space, then delete all the blank space	House price spatial data
saonpaon1	Combine SAON and PAON with a blank space, then delete all the blank space	House price spatial data
saonpaon2	Combine SAON and PAON with a comma and a blank space, then delete all the blank space and all the blank space	House price spatial data
saonpaon3	Combine SAON and PAON with a comma and a blank space	House price spatial data
paonstreetlo	Combine PAON and street with a comma and a blank space, then combine locality with a comma and a blank space, then delete all the blank space	House price spatial data
paonstreetnlo	Combine paonn and streetn with a comma and a blank space, then combine localityn with a comma and a blank space, then delete all the blank space	House price spatial data
paonstreetlo1	Combine PAON and street with a blank space, then combine locality with a comma and a blank space,	House price spatial data

Variable	Create method	Dataset
	then delete all the blank space	
paonstreetnlo1	Combine paonn and streetn with a blank space, then combine localityn with a comma and a blank space, then delete all the blank space	House price spatial data
paonstreetlo2	Combine PAON and street with a blank space, then combine locality with a blank space, then delete all the blank space and comma	House price spatial data
paonstreetn	Combine PAON and streetn with a comma and a blank space, then delete all the blank space	House price spatial data
paon66	Combine paon61 and paon62 with a comma and a blank space, then delete the blank space	House price spatial data
paon65streetlo	Combine paon65 and street with a comma and a blank space, then combine locality with a comma and a blank space, then delete all the blank space	House price spatial data
paon65streetnlo	Combine paon65n and streetn with a comma and a blank space, then combine localityn with a comma and a blank space, then delete all the blank space	House price spatial data
paon65streetlo1	Combine paon65 and street with a blank space, then combine locality with a blank space, then delete all the blank space and comma	House price spatial data
paon61streetlo	Combine paon61 and street with a comma and a blank space, then combine locality with a comma and a blank space, then delete all the blank space	House price spatial data
paon61streetlo1	Combine paon61 and street with a blank space, then combine locality with a blank space, then delete all the blank space and comma	House price spatial data
paon61lo	Combine paon61 and locality with a comma and a blank space, then delete all the blank space	House price spatial data
paon61street	Combine paon61 and street with a blank space, then delete all the blank space and comma	House price spatial data
paon65street	Combine paon65 and street with a blank space, then delete all the blank space and comma	House price spatial data
paon66streetlo	Combine paon62 and paon61 with a blank space, then combine street with a blank space, then combine locality with a blank space, then delete all the comma and blank space	House price spatial data
paon65streetlo	Combine paon65 and street with a comma and a blank space, then combine locality with a comma and a blank space, then delete all the comma and blank space	House price spatial data
paon61new	Combine 'THE' and paon61 with a blank space	House price spatial data

Variable	Create method	Dataset
paonstreetlo3	Combine PAON and street with a comma and a blank space, then combine locality with a comma and a blank space, then delete all the blank space and comma	House price spatial data
paonstreet	Combine PAON and street with a comma and a blank space, then delete all the comma and blank space	House price spatial data
paonstreetn1	Combine PAON and streetn1 with a comma and a blank space, then delete all the comma and all the blank space	House price spatial data
paonstreet1	Combine PAON and street with a comma and a blank space, then delete all blank space	House price spatial data
paonstreet2	Combine PAON and street with a blank space, then delete all blank space	House price spatial data
paon62streetlo	Combine paon62 and street with a comma and a blank space, then combine locality with a comma and a blank space, then delete all the blank space	House price spatial data
paon62streetlo1	Combine paon62 and street with a blank space, then combine locality with a blank space, then delete all the blank space and comma	House price spatial data
paonflat	Combine ‘FLAT’ string and PAON with a blank space	House price spatial data
paonfstreet	Combine paonflat with street with a comma and a blank space, then delete all the blank space	House price spatial data
paonap	Combine ‘APARTMENT’ string and PAON with a blank space	House price spatial data
paonapstreet	Combine paonap with street with a comma and a blank space, then delete all the blank space	House price spatial data
paonfstreet1	Combine paonflat with street with a blank space, then delete all the blank space	House price spatial data
paonfstreetn5	Combine paonflat with streetn5 with a blank space, then delete all the blank space	House price spatial data
paonstreet3	Combine PAON and street with a blank space, then delete all blank space and comma	House price spatial data
paonapstreet1	Combine paonap with street with a blank space, then delete all the blank space	House price spatial data
paonapstreet2	Combine paonap with street with a blank space, then delete all the blank space and comma	House price spatial data
paonapstreetn5	Combine paonap with streetn5 with a blank space, then delete all the blank space	House price spatial data
paonstreet4	Replace ‘FLAT’ to ‘APARTMENT’ in paonstreet3	House price spatial data

Variable	Create method	Dataset
paonfl1	Combine 'FLAT,' string and PAON with a blank space	House price spatial data
paonf1streetn5	Combine paonfl1 with streetn5 with a comma and a blank space, then delete all the blank space	House price spatial data
paonfstreetn6	Combine paonflat with streetn5 with a comma and a blank space, then delete all the blank space	House price spatial data
flpaon3streetn5	Combine 'FLAT ' string and paon with a blank space, then combine with streetn5 with a blank space then delete all the blank space and '-'	House price spatial data
saonpaon65street	Combine SAON and paon65 with a comma and a blank space, then combine street with a comma and a blank space, then delete all the blank space	House price spatial data
saonpaon62streetn2	Combine SAON and paon62 with a comma and a blank space, then combine streetn with a blank space, then delete all the blank space	House price spatial data
saonpaon61street	Combine SAON and paon61 with a blank space, then combine street with a comma and a blank space, then delete all the blank space and comma	House price spatial data
saonpaon62streetn	Combine SAON and paon62 with a blank space, then combine streetn with a blank space, then delete all the blank space	House price spatial data
saonpaonn	Combine saonn and paonn with a comma and a blank space, then delete all the blank space	House price spatial data
saon2street	Combine saon2 and street with a comma and a blank space, then delete all the blank space	House price spatial data
saon2paon61street	Combine saon2 and paon61 with a blank space, then combine street with a comma and blank space, then delete all the blank space.	House price spatial data
flsaonpaonstreet0	Combine flsaon and PAON with a comma and a blank space and then combine street with a comma and a blank space	House price spatial data
flsaonpaon1	Combine flsaon and PAON with a blank space, then delete all the blank space	House price spatial data
flsaonpaon2	Combine flsaon and PAON with a comma and a blank space, then delete all the blank space	House price spatial data
flsaon	For the SAON start with number string, combine 'FLAT' string with SAON with a blank space	House price spatial data
flsaon1	For the SAON start with number string, combine 'FLAT' string with saonn with a blank space	House price spatial data
flsaon3	combine 'FLAT' string with SAON with a blank space	House price spatial data

Variable	Create method	Dataset
flsaon1paonstreetn2	Combine flsaon1 with paonn with a comma and a blank space, then combine the streetn2 with a comma and a blank space, then delete all the blank space	House price spatial data
flsaonpaonstreet1	Combine flsaon with PAON with a blank space, then combine the street with a blank space, then delete all the blank space and comma	House price spatial data
flsaonpaon62street1	Combine flsaon and paon62 with a blank space, then combine street with a blank space, then delete all the blank space and comma	House price spatial data
fldsaonpaonstreet1	Combine fldsaon and PAON with a blank space, then combine street with a blank space, then delete all the blank space and comma	House price spatial data
saon7paonstreet1	Combine saon7 and PAON with a comma and a blank space, then combine street with a blank space, then delete all the blank space	House price spatial data
saon7paonstreet2	Combine saon7 and PAON with a blank space, then combine street with a blank space, then delete all the blank space and comma	House price spatial data
apsaon	For SAON starts with number string, combine 'APARTMENT' string with SAON with a blank space	House price spatial data
apsaonpaonstreet1	Combine apsaon and PAON with a blank space, then combine street with a blank space, then delete all the blank space and comma	House price spatial data
saon7paonstreetn	Combine saon71 and paonn with a comma and a blank space, then combine street with a blank space, then delete all the blank space	House price spatial data
saon7paonn	Combine saon7 and paonn with a comma and a blank space, then delete all the blank space	House price spatial data
saon7paon	Combine saon7 and PAON with a comma and a blank space, then delete all the blank space	House price spatial data
saon4paonstreetn	Combine saonn4 and paonn with a comma and a blank space, then combine streetn with a blank space, then delete all the blank space	House price spatial data
saon4paonstreetn1	Combine saonn4 and paonn with a blank space, then combine streetn with a comma and a blank space, then delete all the blank space	House price spatial data
apsaonpaon6streetn	Combine apsaon and paon62 with a comma and a blank space, then combine streetn with a blank space, then delete all the blank space	House price spatial data
flsaonpaonstreetn	Combine flsaon and PAON with a comma and a blank space, then combine with streetn with a blank space, then delete all the blank space	House price spatial data

Variable	Create method	Dataset
saon4paonstreetn3	Combine saonn4 and paonn with a blank space, then combine streetn with a blank space, then delete all the blank space	House price spatial data
saon4paonstreetn4	Combine saonn4 and paonn with a comma and a blank space, then combine streetn with a comma and a blank space, then delete all the blank space	House price spatial data
saon1paonstreetn	combine saon1 and paonn with a blank space, then combine streetn with a comma and a blank space, then delete all the blank space	House price spatial data
saon1paonstreetn1	Combine saon1 and paonn with a comma and a blank space, then combine streetn with a comma and a blank space, then delete all the blank space	House price spatial data
saon1paonstreetn2	Combine saon1 and paonn with a blank space, then combine streetn with a blank space, then delete all the blank space and comma	House price spatial data
saon2paonstreetn3	Combine saonn2 and paonn with a blank space, then combine streetn with a blank space, then delete all the blank space	House price spatial data
saon2paonstreetn2	Combine saonn2 and paonn with a blank space, then combine streetn with a comma and a blank space, then delete all the blank space	House price spatial data
saonn2paonn1	Combine saonn2 and paonn with a blank space, then delete all the blank space	House price spatial data
saonpaon62street	Combine SAON and paon62 with a comma and a blank space, then combine street with a blank space, then delete all the blank space	House price spatial data
saon2paonstreetn	Combine saonn2 and paonn with a comma and a blank space, then combine street with a blank space, then delete all the blank space	House price spatial data
saonn3paonnstreet	Combine saonn3 and paonn with a comma and a blank space, then combine street with a blank space, then delete all the blank space	House price spatial data
saonn2paonn1streetn	Combine saonn2 and paonn with a comma and a blank space, then combine street with a blank space, then delete all the blank space	House price spatial data
saonpaon62streetn1	Combine SAON and paon62 with a comma and a blank space, then combine streetn with a comma and a blank space, then delete all the blank space	House price spatial data
saon1paonstreet6n	Combine saon1 and paon62 with a comma and a blank space, then combine streetn with a comma and a blank space, then delete all the blank space	House price spatial data
saon1paonstreet6n1	Combine saon1 and paon62 with a comma and a blank space, then combine street with a blank space,	House price spatial data

Variable	Create method	Dataset
	then delete all the blank space	
saon2paonstreetn4	Combine saonn2 and paonn with a comma and a blank space, then combine streetn with a comma and a blank space, then delete all the blank space	House price spatial data
saon5paonstreetn1	Combine saonn5 and paonn with a blank space, then combine streetn with a comma and a blank space, then delete all the blank space	House price spatial data
paonsaon2streetn	Combine paonn and saonn2 with a blank space, then combine streetn with a blank space, then delete all the blank space	House price spatial data
paon62saonpstreet	Combine paon62 and saon with a blank space, then combine paon61 with a blank space and then combine street with a comma and a blank space, then delete all the blank space and comma	House price spatial data
saonpaon66street	Combine saonn and paon62 with a comma and a blank space, then combine paon61 with a blank space, then combine street with a blank space, then delete all the blank space and comma	House price spatial data
saon1paonstreetn3	Combine saon1 and PAON with a comma and a blank space, then combine streetn with a blank space, then delete all the blank space	House price spatial data
saon1paonstreet	Combine sao1 and PAON with a comma and a blank space, then combine street with a blank space, then delete all the blank space	House price spatial data
saon2paonlo	Combine saon2 and PAON with a blank space, then combine locality with a comma and a blank space, then delete all the blank space	House price spatial data
saon1paon	Combine sao1 and PAON with a comma and a blank space, then delete all the blank space	House price spatial data
saon1paon61street	Combine sao1 and paon61 with a blank space, then combine street with a comma and a blank space, then delete all the blank space	House price spatial data
saon1paon1	Combine sao1 and PAON with a blank space, then delete all the blank space	House price spatial data
psaonpaonstreet	Combine paon64 and SAON, then combine paon641 with a blank space, then combine street with a comma and a blank space, then delete all the blank space and comma	House price spatial data
saon2paon62street	Combine saon2 and paon62 with a comma and a blank space, then combine street with a comma and a blank space, then delete all the blank space	House price spatial data

Variable	Create method	Dataset
saon2paonstreet	Combine saon2 and PAON with a blank space, then combine street with a comma and a blank space, then delete all the blank space	House price spatial data
flsaonpaonstreet	Combine flsaon with PAON with a comma and a blank space, then combine the street with a comma and a blank space, then delete all the blank space and comma	House price spatial data
psaon8street	Combine PAON and fldsaon1, then combine street with a blank space then delete al the blank space and comma	House price spatial data
saonstreet	Combine SAON and street with a comma and a blank space, then delete all the blank space	House price spatial data
saonstreet1	Combine SAON and street with a blank space, then delete all the blank space and comma	House price spatial data
saonstreet2	Combine SAON and street with a comma and a blank space, then delete all the blank space and comma	House price spatial data
saonstreet3	Combine SAON and street with a blank space, then delete all the blank space	House price spatial data
saonstreetlo	Combine SAON and street with a comma and a blank space, then combine with locality with a comma and a blank space, then delete all the blank space	House price spatial data
unsaonpaonstreet2	Combine 'UNIT' string with SAON with a blank space, then combing PAON with a blank space, then combine with street with a comma and a blank space and the delete all the blank space.	House price spatial data
flsaonpaonstreet2	Combine flsaon2 with PAON with a blank space, then combine the street with a comma and a blank space, then delete all the blank space	House price spatial data
saon7paon6street	Combine saon7 and paon62 with a comma and a blank space, then combine street with a blank space, then delete all the blank space	House price spatial data
saon8paonstreet2	Combine saon8 and PAON with a blank space, then combine street with a comma and a blank space, then delete all the blank space	House price spatial data
paonlo	For PAON start with number string, combine PAON and locality with a comma and a blank space, then delete all the blank space.	House price spatial data
flsaonpaonstreet3	Combine flsaon with PAON with a blank space, then combine the street with a comma and a blank space, then delete all the blank space and comma	House price spatial data
saonpaon62steet	Combine SAON and paon62 with a comma and a blank space, then combine street with a comma and blank space, then delete all the blank space	House price spatial data

Variable	Create method	Dataset
flsaonpaon61street	Combine flsaon with paon61 with a blank space, then combine the street with a comma and a blank space, then delete all the blank space and comma	House price spatial data
flsaonpaon61street1	Combine flsaon with paon61 with a blank space, then combine the street with a comma and a blank space, then delete all the blank space	House price spatial data
saon4paonstreet	Combine saon4 with PAON with a blank space, then combine the street with a blank space, then delete all the blank space	House price spatial data
saonpaon61street1	Combine SAON and paon61 with a blank space, then combine street with a comma and a blank space, then delete all the blank space	House price spatial data
flsaonpaonstreet5	Combine flsaon with PAON with a comma and a blank space, then combine the street with a comma and a blank space, then delete all the blank space	House price spatial data
paonsaonstreet	Combine PAON and SAON, then combine street with a comma and a blank space, then delete all the blank space	House price spatial data
saonpaon61	Combine SAON and paon61with a comma and a blank space, then delete all the blank space	House price spatial data
paonsaonstreet1	Combine PAON and SAON with a comma and a blank space, then combine street with a comma and a blank space, then delete all the blank space	House price spatial data
apsaonpaon	Combine apsaon and PAON with a blank space, then delete all the blank space	House price spatial data
saon1paon62street	Combine sao1 and paon62 with a comma and a blank space, then combine street with a comma and a blank space, then delete all the blank space	House price spatial data
apsaonpaon62street1	Combine apsaon and paon62 with a comma and a blank space, then combine street with a blank space, then delete all the blank space	House price spatial data
saon2paonstreet1	Combine saon2 and PAON with a blank space, then combine street with a comma and a blank space	House price spatial data
apsaonpaonstreet2	Combine apsaon and PAON with a blank space, then combine street with a comma and a blank space, then delete all the blank space	House price spatial data
psaonpstreet	Combine paon6164 and SAON, then combine paon6163 with a blank space, then combine paon62 with a comma and hen combine street with a comma and a blank space and delete all the blank space	House price spatial data
saonpaonstreet11	Combine SAON and paon11 with a blank space, then combine street with a blank space, then delete all the blank space	House price spatial data

Variable	Create method	Dataset
saonpaon61streetn	Combine saonn and paonn61 with a comma and a blank space and then combine streetn with a comma and a blank space, and then delete all the blank space	House price spatial data
saonpaon65street1	Combine SAON and paon65 with a comma and a blank space, then combine street with a blank space, then delete all the blank space	House price spatial data

Appendix A3

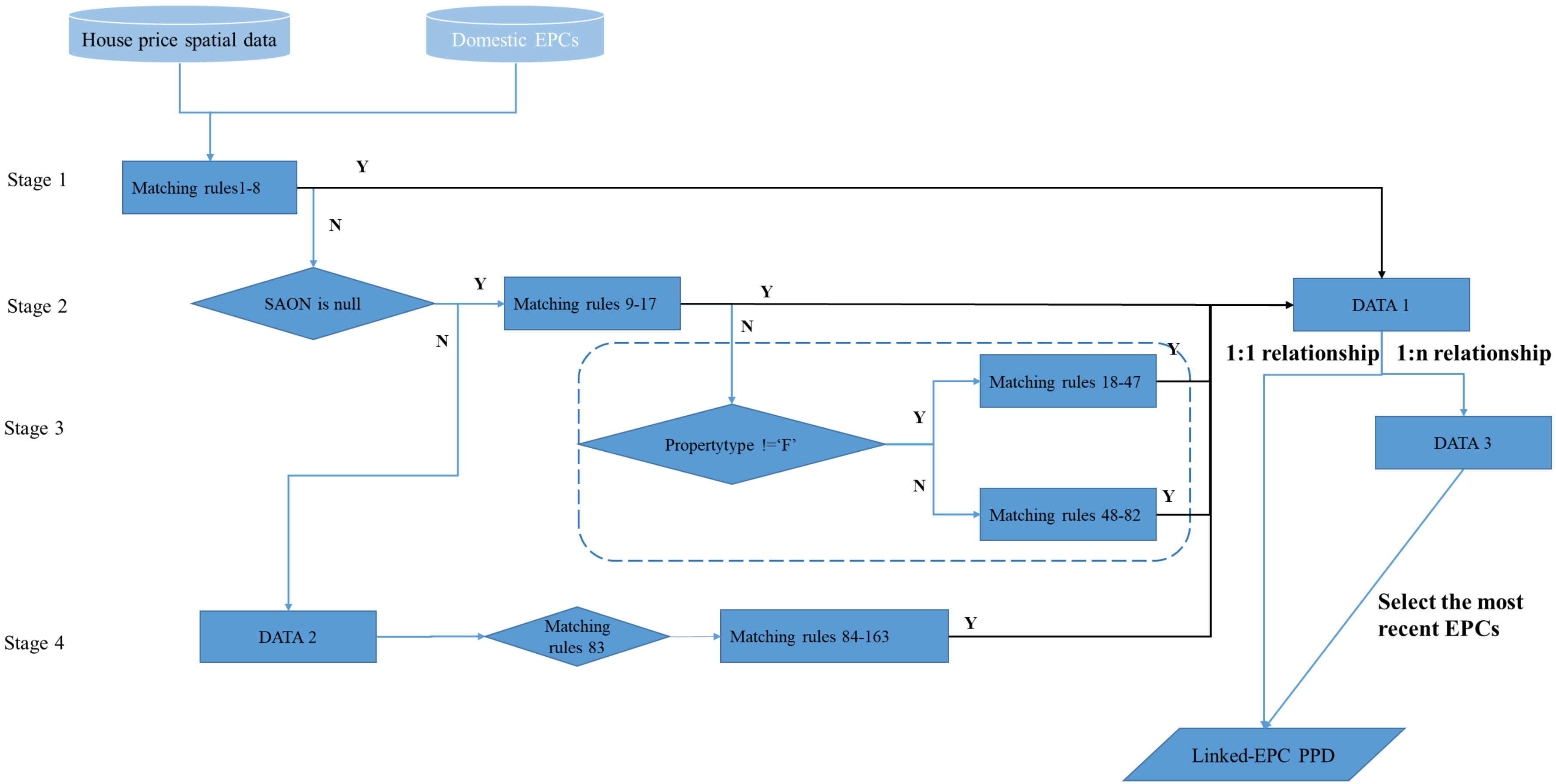


Figure A3 Master workflow of the 4 stages data linkage between house price spatial data and Domestic EPC

Figure A3 demonstrates the data linkage workflow between Domestic EPCs and house price spatial data. Each stage contains more than one match rule. Details of the match rules for each Stage are listed in Table A3. In the Domestic EPCs, each record is created using a unique identifier with names of epcid. Each transaction in house price spatial data has a unique identifier named transactionid. The whole matching process between these two datasets is divided into four Stages. Take Stage one as an example of the matching process. All the matches are based on a "temple address strings" (i.e. postcode+saonpaonstreet) which is the combination of postcode and address variables. When Domestic EPCs and house price spatial data are put into the matching process, the process starts to link house price spatial data (transactionid) with epcid basing on the "temple address strings". For example, it tests whether postcode+saonpaonstreet in house price spatial data is equal to any postcode +ADDRE in Domestic EPCs. If the result shows yes and the epcid will direct link with transactionid and restore in Data 1, otherwise the transaction records will move to the other matching rules within the same Stage to conduct further matching tests, For the transactions that cannot be matching in Stage 1, they will move to Stage 2 to do the further matching tests. All the successfully linked transactions in Stage 1 are stored in DATA 1. However, in the real world, one property could have more than one Domestic EPCs in this matching process. The transaction property with only one successfully linked EPC will direct stored in linked-EPC PPD, transaction property with successful links to more than one EPC will be stored in DATA 3. A new function will be conducted to select all the Domestic EPCs for which total floor area is not null or 0 and will then keep the EPC's inspection date or lodgement date which is closest to the transaction date in the house price data. This result will then be stored in linked-EPC PPD. Stages 2 to 4 follow a similar process to

Stage 1. Finally, linked-EPC PPD is the data linkage result. These data linkage results will firstly join back to Domestic EPCs according to the same epcid, then join with house price spatial data according to the transactionid. The data linkage process is conducted in R Studio.

Table A3 Details of matching rules in 4 stages[37]

Stage No.	Match rules No.	Match rules
Stage 1	1	(saonpaonstreet OR saonpaonstreet1 OR saonpaonstreet2 OR saonpaonlo OR saonpaonstreetlo OR saonpaonstreet3 OR saonpaon1) = ADDRE;
	2	(saopaonstreetn ORsaonpaonstreetn1 OR saopaonstreetn2 OR saonpaonlon OR saonpaonstreetnlo or saonpaonstreetn3) = ADDC
	3	(saonpaonstreet OR saonpaonstreet1 OR saonpaonstreet2 OR saonpaonstreet3 OR saonpaonstreetlo or saonpaonlo) = ADD12;
	4	(saopaonstreetn OR saonpaonstreetn1 OR saopaonstreetn2 OR saonpaonlon) = ADD12C;
	5	saonpaonlon = ADDCC;
	6	saonpaonstreetn3 = ADD12C1;
	7	saonpaonstreet31 = ADDREC
	8	saonpaonstreetn31 = ADDC3
Stage 2	9	(paonstreetlo OR paonstreetlo1) = ADDRE;
	10	(paonstreetnlo OR paonstreetnlo1) = ADDC;
	11	(paonstreetlo OR paonstreetlo1) = ADD12
	12	(paonstreetnlo OR paonstreetnlo1) = ADD12C
	13	paonstreetlo2= ADD12C2
	14	paonstreetlo2= ADDREC
	15	paonstreetn=ADD12C3
	16	street is null and paonn3 =ADD1CC

[37] In this table, all the address fields in house price spatial data is written in small letters and the address variable in the Domestic EPCs is written capital letters.

Stage No.	Match rules No.	Match rules
	17	For the PAON contain comma, then paon66=ADD1CC
Stage 3	18	paon65streetlo=ADDRE
	19	paon65streetlo=ADD12
	20	paon65streetnlo =ADDCC
	21	(paon65streetlo1 OR paon61streetlo1)=ADDREC
	22	paon61streetlo=ADDC
	23	(paon61streetlo1 OR paon65street) = ADDC3
	24	paon61streetlo1= ADD12C1
	25	paon61lo= ADD12C
	26	paon61street= ADD12C1
	27	paon61street= ADD13C1
	28	paon65street= ADD1C2
	29	paon66streetlo=ADDCCC
	30	paon66streetlo =ADD12C3
	31	For the propertytype in EPCs is not Flat or Maisonette, paon65streetlo1 =ADD23C1
	32	For the propertytype in EPCs is not Flat or Maisonette, paon61new=ADD1
	33	paonstreetlo3= ADD12new
	34	paonstreetlo3= ADD13C1
	35	paonstreetlo3 = ADD13C2
	36	paonstreet= ADD1C3
	37	PAON=ADD1
	38	paonstreetlo3 =ADD662

Stage No.	Match rules No.	Match rules
	39	paonstreet= ADD67
	40	For the street is not null and the propertytype in EPCs is not Flat or Maisonette , paonstreet= ADDSP12;
	41	paonstreetn1=ADD1C4
	42	For the propertytype in EPCs is not Flat or Maisonette, paonstreet=ADDU
	43	paonstreet1=ADD68
	44	paonstreet1=ADD69
	45	For the address are written differently(e.g "WOODLANDS PARK" vs "WOODLAND PARK"), (paonstreet1 OR paonstreet2) =ADD1C5
	46	For the address are written differently and the propertytype in EPCs is not Flat or Maisonette, paonn2=ADD1C6
	47	For the ADD did not have 'number - number' (i.e 3-5), paonstreet3=ADDCCC
	48	For the paon61 did not contain ‘FLAT’ string and ‘FLOOR’ string, then paon62streetlo= ADDRE;
	49	For the paon61 does not contain FLAT’ string and ‘FLOOR’ string and also not start with number, then paon62streetlo=ADD12
	50	paon65streetnlo=ADDCC;
	51	For property type in EPC is Flat/Maisonette, paon62streetlo1=ADDREC;
	52	For property type in EPC is Flat/Maisonette, (paon61streetlo OR paon61streetlo1)=ADDC;
	53	paon61streetlo1=ADDC3
	54	paon61streetlo1=ADD12C1
	55	For property type in EPC is Flat/Maisonette, paon61street= ADD13C1
	56	paon66streetlo= ADDCCC
	57	paon66streetlo =ADD12C3
	58	paonfstreet= ADD12
	59	(paonfstreet OR paonapstreet OR paonfstreet1)= ADDRE

Stage No.	Match rules No.	Match rules
	60	(paonstreet OR paonstreetn1)= ADD1C7
	61	paonstreetn1= ADD1C8
	62	For the address words written different, (paonstreet1 OR paonstreet2)=ADD1C5; PAON=ADD1df1; **paonn2=ADD1du**; paon61c=ADD1C9; paonfstreetn5=ADD1C3;
	63	For PAON starts with number string, (paonfstreetn5=ADD1C9
	64	For PAON starts with number string, (paonfstreetn5 OR paonstreet1)=ADD1C;
	65	For property type in EPC is Flat/Maisonette, paonstreet3=ADD1632;
	66	For PAON starts with number string and add2 in EPCs not starts with number string, paonapstreet1=ADD12C2
	67	For PAON starts with number string, paonapstreetn5=ADD12C1
	68	(paonn2 OR paonstreet4)=ADDC3
	69	paonstreet3=flADD
	70	paonn2=ADD2611
	71	paonstreet3=flADD13
	72	paonstreet3=ADD13C2
	73	For PAON starts with number string, paonfstreetn5=ADD1C2
	74	(paonn2 OR paonstreet2)=ADD1C2
	75	paonfstreetn6= ADD12C
	76	paonapstreet2=ADD12C2
	77	paonf1streetn5=ADD12C;
	78	For the add in EPC is not start with 'number string, number stirng' pattern, for the PAON start with number but do not contain '-' in number string. flpaon3streetn5=ADDC10
	79	paonstreet2=ADD5

Stage No.	Match rules No.	Match rules
	80	paonstreet2=apADD1
	81	paonapstreet2=ADD13C2
	82	paonstreet3=ADDr66
Stage 4	83	Correct the address components in EPC basing on address components in PPD (e.g "GREENFELL COURT" to "GRENFELL COURT")
	84	saonpaonstreet2=ADDRE
	85	saonpaonstreet2=ADD12
	86	saonpaonstreetn=ADDC
	87	saonpaon65street=ADD12C;
	88	saonpaon62streetn2=ADD13C
	89	saonpaonstreetn=ADD6
	90	saonpaonstreetn=ADDCC
	91	saonpaon61street=ADD12C2
	92	saonpaon61street=ADDREC
	93	saonpaon62streetn=ADD7
	94	saonpaonstreet1=ADD13C2
	95	saonpaon1=ADD1C9
	96	saonpaonn=ADDC4
	97	paonstreetn=ADDC4 and saon='FLAT';
	98	For the property type is not Flats/Maisonettes;paonstreetn=ADDC4 , then delete keep the successful linkage whose properety type in EPC is not Flat or Maisonette

Stage No.	Match rules No.	Match rules
	99	For property type in EPC is Flat/Maisonette and for propertytype in house price data is ‘F’ :saon2paon61street= ADDCC; fldsaonpaonstreet1=ADDREC
	100	saonpaonn=ADD12C;
	101	paonstreetn=ADDC4 and saon did not contain ‘FLOOR’,’UPPER’,’BASEMENT’, ‘LOWER’ ,‘FLAT ’ or any number string
	102	For propertytype in house price data is ‘F’, (flsaonpaonstreet0=ADD; flsaon1paonstreetn2=ADDCC; flsaonpaonstreet1= ADDREC; flsaonpaon62street1 = ADDREC; saon7paonstreet1=ADDRE; saon7paonstreet2=ADDREC;saon7paonstreet2=ADD12C2 ; apsaonpaonstreet1=ADD12C2)
	103	For propertytype in house price data is ‘F’ and SAON start with number string, apsaonpaonstreet1=ADDREC
	104	saon7paonstreetn=ADDC4
	105	saon7paonn=ADD12C4
	106	saon4paonstreetn=ADDC4
	107	apsaonpaon6streetn=ADDC4
	108	For propertytype in house price data is ‘F’, flsaonpaonstreetn=ADDC4
	109	For the PAON start with number string , saon4paonstreetn3=ADDC5
	110	saon4paonstreetn4=ADD12C
	111	For property type in EPC is Flat/Maisonette and for propertytype in house price data is ‘F’, saon4paonstreetn1=ADD12C
	112	For propertytype in house price data is ‘F’: saon1paonstreetn=ADDC; saon1paonstreetn=ADD12C; saon1paonstreetn1=ADDC ;saon1paonstreetn1=ADD12C; saon1paonstreetn2=ADDC3; saon1paonstreetn2=ADD12C1;
	113	For property type in EPC is Flat/Maisonette and for propertytype in house price data is ‘F’, saon2paon61street=ADD12C
	114	For propertytype in house price data is ‘F’: saon2paonstreetn3=ADDC; saon2paonstreetn3=ADD12C
	115	For propertytype in house price data is ‘F’ and PAON start with number string: saon2paonstreetn2=ADDC; saon2paonstreetn2=ADD12C

Stage No.	Match rules No.	Match rules
	116	saonn2paonn1=ADDC
	117	saonpaon62street=ADD12C
	118	saon1paonstreet6n1=ADD12C
	119	For property type in EPC is Flat/Maisonette and for propertytype in house price data is ‘F’, saon2paonstreetn=ADD12C
	120	saonn3paonnstreet=ADD13C
	121	saonn2paonn1streetn=ADDC
	122	saonpaon62streetn1=ADDC
	123	saon1paonstreet6n=ADD12C
	124	For propertytype in house price data is ‘F’: paon62saonpstreet= ADDREC; saon2paonstreetn4=ADDC; saon2paonstreetn4=ADD12C
	125	For propertytype in house price data is ‘F’: saon2paonstreetn4=ADD1num2 and ADD1 in EPC does not contain a character pattern that consist of number strings with a character
	126	For propertytype in house price data is ‘F’ and SAON contain ‘APARTMENT ’ string: saon5paonstreetn1=ADDC; paonsaon2streetn=ADD1C; saon2paonstreetn2=ADD13C; saonpaon61streetn=ADDC
	127	saonpaon66street=ADDC6
	128	For propertytype in house price data is ‘F’: saon1paonstreetn3=ADD12C
	129	For property type in EPC is Flat/Maisonette and for propertytype in house price data is ‘F’: saon2street=ADDC; saon2paonlo=ADDRE; saon2paonstreet=ADD12
	130	saon1paonstreet=ADDRE
	131	saon1paon=ADD12
	132	For propertytype in house price data is ‘F’: saon1paon61street=ADD12; saon1paon1=ADD1; saon1paonstreetn2=ADD12C2; psaonpaonstreet=ADDRE
	133	For property type in EPC is Flat/Maisonette and for propertytype in house price data is ‘F’: saon2paon62street=ADD12

Stage No.	Match rules No.	Match rules
	134	For add2 contain 'number - number' character pattern , saon2paonstreet=ADD1262
	135	saonpaonstreetn2=ADD7
	136	For property type in EPC is Flat/Maisonette and for propertytype in house price data is ‘F’: flsaonpaonstreet=add1f61f2; psaon8street=ADDREC; saonpaonstreet1=add12643
	137	For propertytype is not F,saonstreet=ADDRE
	138	saonstreetlo= ADDRE
	139	For SOAN starts with number string, unsaonpaonstreet2=ADDRE
	140	For propertytype in house price data is ‘F’: flsaonpaonstreet2=ADD8; saon7paon6street=ADDRE; saon7paon6street=ADD12
	141	For property type in EPC is Flat/Maisonette and for propertytype in house price data is ‘F’: flsaonpaon1=ADD1C9; saonpaon1=fladd; saonpaon1=fladd1c; saonpaonstreet3=fladd
	142	saon8paonstreet2=ADDRE
	143	For property type in EPC is Flat/Maisonette and for propertytype in house price data is ‘F’: saonpaonstreet2=fladd
	144	PAON start with number string, paonlo=ADD12;
	145	For propertytype in house price data is ‘F’: saonpaonstreet1=adddap; saonpaon2=fladdc; saonpaonstreet11=ADD12; saonpaon61street=ADD1262C and paon62 contain ‘-’ string; saonpaon61street=ADD1262C and add261 contain ‘-’ string; flsaonpaonstreet3=ADD12C5; flsaonpaonstreet3=ADD12C1; saonpaon62steet=ADDC7
	146	For property type in EPC is Flat/Maisonette and for propertytype in house price data is ‘F’: saonpaon61street=fladdc; saonpaonstreet5 =apadd1632; saonstreet1=ADD1C7; saonpaonstreet1=add1f61f2; saonstreet2=ADD1264;flsaonpaon61street=ADDREC;saon4paonstreet=ADD12
	147	For propertytype in house price data is ‘F’: saonpaon61street1=ADD1262; flsaonpaon2=ADDRE
	148	saonpaon3=ADD1

Stage No.	Match rules No.	Match rules
	149	For propertytype in house price data is ‘F’: saonstreet3=ADDC; flsaonpaon2=ADD12; flsaonpaonstreet5=ADD1262; PAON did not contain number string and comma, saonstreet=ADD1264
	150	paonsaonstreet=ADDRE
	151	For propertytype in house price data is ‘F’: saonpaon61=ADD12; saon7paon=ADD12; paonsaonstreet1=ADD12; flsaonpaon61street1=ADD12
	152	For propertytype in house price data is ‘F’ : apsaonpaon=ADD12C2; apsaonpaon62street1=ADDC8
	153	saon1paon62street=ADD12
	154	For propertytype in house price data is ‘F’ and PAON does not start with number string, For property type in EPC is Flat/Maisonette, saonstreet=ADDC5
	155	saonpaonstreet2=ADDRE
	156	For propertytype in house price data is ‘F’: saon2paonstreet1=ADDC9; apsaonpaonstreet2=ADD1262cc; psaonpstreet=ADDRE
	157	saonpaon65street1=ADD12C
	158	For propertytype in house price data is ‘F’: saon2paonstreetn3=ADDC
	159	saonpaonn=ADD12C
	160	saon1paonstreetn1=ADDC
	161	For property type in EPC is Flat/Maisonette and ADD in EPC does not contian 'number -number' pattern and propertytype in house price data is ‘F’, saon4paonstreetn1=ADDC4
	162	For propertytype in house price data is ‘F’: saon1paonstreetn=ADDC4
	163	saonpaonlon=ADDC4

Appendix A4

Table A4 Unable matched in method 2 due to the missing postcode in EPCs by LA

LA	House price spatial data	Unable matched	Proportion	Region
City of London	2218	189	8.52%	London
Westminster	27254	794	2.91%	London
Salford	26909	772	2.87%	North West
Isles of Scilly	125	3	2.40%	South West
Tower Hamlets	33165	647	1.95%	London
Newcastle upon Tyne	26176	483	1.85%	North East
Liverpool	39210	630	1.61%	North West
Manchester	47723	753	1.58%	North West
Rochdale	17689	251	1.42%	North West
St. Helens	14962	212	1.42%	North West
West Devon	7008	94	1.34%	South West
Brent	20639	271	1.31%	London
Newham	18715	229	1.22%	London
Hackney	20172	243	1.20%	London
Crawley	11311	132	1.17%	South East
Blaby	12020	139	1.16%	East Midlands
Ashford	15818	177	1.12%	South East
North Warwickshire	6307	66	1.05%	West Midlands
Watford	12014	125	1.04%	East of England
Halton	10320	104	1.01%	North West
Redditch	9015	88	0.98%	West Midlands
Vale of White Horse	15374	150	0.98%	South East
Horsham	19208	184	0.96%	South East
Hartlepool	8573	82	0.96%	North East
Selby	10778	102	0.95%	Yorkshire and The Humber
West Oxfordshire	13163	124	0.94%	South East

LA	House price spatial data	Unable matched	Proportion	Region
Cheltenham	17099	156	0.91%	South West
South Tyneside	12482	111	0.89%	North East
Cambridge	13664	121	0.89%	East of England
West Berkshire	18949	167	0.88%	South East
Melton	6150	54	0.88%	East Midlands
Camden	20737	181	0.87%	London
Trafford	26363	228	0.86%	North West
Hertsmere	11922	101	0.85%	East of England
Test Valley	14951	126	0.84%	South East
Peterborough	21360	180	0.84%	East of England
East Dorset	12473	105	0.84%	South West
Erewash	13667	115	0.84%	East Midlands
Barnsley	23505	196	0.83%	Yorkshire and The Humber
Daventry	10367	86	0.83%	East Midlands
Wandsworth	42603	350	0.82%	London
Blackburn with Darwen	11938	97	0.81%	North West
Purbeck	5588	45	0.81%	South West
West Dorset	14339	115	0.80%	South West
Bromsgrove	11174	89	0.80%	West Midlands
West Lancashire	9433	75	0.80%	North West
Warwick	18482	146	0.79%	West Midlands
Middlesbrough	12441	98	0.79%	North East
Ealing	26603	209	0.79%	London
Hounslow	21980	172	0.78%	London
Nottingham	26340	204	0.77%	East Midlands
Rugby	12921	100	0.77%	West Midlands
Wokingham	20556	159	0.77%	South East
Maldon	7503	58	0.77%	East of England
Hyndburn	8286	63	0.76%	North West
Harlow	8699	66	0.76%	East of England

LA	House price spatial data	Unable matched	Proportion	Region
South Northamptonshire	11547	87	0.75%	East Midlands
Maidstone	19936	150	0.75%	South East
Wellingborough	8409	63	0.75%	East Midlands
Cotswold	11487	85	0.74%	South West
Hambleton	9773	72	0.74%	Yorkshire and The Humber
Stevenage	9095	67	0.74%	East of England
Colchester	26020	191	0.73%	East of England
Bedford	20766	151	0.73%	East of England
Tonbridge and Malling	16250	118	0.73%	South East
Cornwall	68595	498	0.73%	South West
Huntingdonshire	23966	173	0.72%	East of England
Barnet	34504	248	0.72%	London
Stockton-on-Tees	20581	147	0.71%	North East
Basildon	20405	145	0.71%	East of England
South Ribble	12693	90	0.71%	North West
Darlington	11282	79	0.70%	North East
Southwark	28234	195	0.69%	London
Leicester	24474	169	0.69%	East Midlands
Bolton	25249	174	0.69%	North West
Bradford	46397	319	0.69%	Yorkshire and The Humber
Bolsover	8513	58	0.68%	East Midlands
Wealden	21302	145	0.68%	South East
East Hertfordshire	19688	134	0.68%	East of England
Calderdale	22405	151	0.67%	Yorkshire and The Humber
Malvern Hills	8462	57	0.67%	West Midlands
Stratford-on-Avon	15938	107	0.67%	West Midlands

LA	House price spatial data	Unable matched	Proportion	Region
Norwich	16912	113	0.67%	East of England
North Hertfordshire	16798	111	0.66%	East of England
Slough	12295	81	0.66%	South East
Lichfield	11304	74	0.65%	West Midlands
Sandwell	22880	149	0.65%	West Midlands
Bassetlaw	11839	77	0.65%	East Midlands
South Hams	11762	76	0.65%	South West
Eden	5456	35	0.64%	North West
Uttlesford	11718	75	0.64%	East of England
Winchester	14869	94	0.63%	South East
Dartford	13631	86	0.63%	South East
Gosport	10472	66	0.63%	South East
Broxtowe	12756	79	0.62%	East Midlands
Chichester	15834	98	0.62%	South East
Wakefield	31928	197	0.62%	Yorkshire and The Humber
Derby	26310	162	0.62%	East Midlands
Harrogate	20403	125	0.61%	Yorkshire and The Humber
Shropshire	32673	199	0.61%	West Midlands
Sefton	25542	155	0.61%	North West
Stafford	14506	88	0.61%	West Midlands
Amber Valley	14721	89	0.60%	East Midlands
Cheshire West and Chester	37100	223	0.60%	North West
Northumberland	33116	199	0.60%	North East
Wychavon	14427	86	0.60%	West Midlands
Bury	19149	114	0.60%	North West
Stockport	32647	193	0.59%	North West
North Devon	12039	71	0.59%	South West
Torbay	19180	113	0.59%	South West
Waverley	16241	95	0.58%	South East

LA	House price spatial data	Unable matched	Proportion	Region
North Tyneside	23460	137	0.58%	North East
Burnley	9442	55	0.58%	North West
Derbyshire Dales	7897	46	0.58%	East Midlands
Kirklees	39670	230	0.58%	Yorkshire and The Humber
Mendip	14538	84	0.58%	South West
Swindon	28678	165	0.58%	South West
Adur	8198	47	0.57%	South East
Northampton	27255	155	0.57%	East Midlands
Wiltshire	59442	338	0.57%	South West
Milton Keynes	33118	188	0.57%	South East
Sunderland	22382	127	0.57%	North East
Bristol, City of	55132	312	0.57%	South West
Leeds	82385	466	0.57%	Yorkshire and The Humber
Preston	13104	74	0.56%	North West
Warrington	22188	125	0.56%	North West
Great Yarmouth	11380	64	0.56%	East of England
Lincoln	11982	67	0.56%	East Midlands
Richmondshire	4709	26	0.55%	Yorkshire and The Humber
Lambeth	31941	176	0.55%	London
Birmingham	86831	478	0.55%	West Midlands
Oldham	17733	97	0.55%	North West
Tamworth	6957	38	0.55%	West Midlands
South Lakeland	13011	71	0.55%	North West
Forest Heath	8995	49	0.54%	East of England
Welwyn Hatfield	11945	65	0.54%	East of England
Runnymede	10485	57	0.54%	South East
Braintree	18580	101	0.54%	East of England
Cannock Chase	9400	51	0.54%	West Midlands
Allerdale	10150	55	0.54%	North West

LA	House price spatial data	Unable matched	Proportion	Region
South Oxfordshire	17385	94	0.54%	South East
Redcar and Cleveland	12990	70	0.54%	North East
Tewkesbury	11337	61	0.54%	South West
Wyre Forest	10244	55	0.54%	West Midlands
Reading	20499	110	0.54%	South East
Babergh	10687	57	0.53%	East of England
Doncaster	27755	147	0.53%	Yorkshire and The Humber
Bath and North East Somerset	22575	119	0.53%	South West
Corby	8161	43	0.53%	East Midlands
Sutton	23180	122	0.53%	London
East Hampshire	15243	80	0.52%	South East
Harrow	20033	105	0.52%	London
Aylesbury Vale	25411	133	0.52%	South East
Canterbury	20267	106	0.52%	South East
Oxford	13613	71	0.52%	South East
Eastbourne	15864	82	0.52%	South East
Mid Sussex	20139	104	0.52%	South East
Boston	7166	37	0.52%	East Midlands
Stroud	14545	75	0.52%	South West
Forest of Dean	8729	45	0.52%	South West
South Derbyshire	12522	64	0.51%	East Midlands
Wigan	29441	149	0.51%	North West
North Dorset	9094	46	0.51%	South West
Telford and Wrekin	17793	90	0.51%	West Midlands
Ribble Valley	6529	33	0.51%	North West
County Durham	50896	257	0.50%	North East
Gateshead	18482	93	0.50%	North East
Taunton Deane	15734	79	0.50%	South West

LA	House price spatial data	Unable matched	Proportion	Region
Exeter	15753	79	0.50%	South West
Oadby and Wigston	5786	29	0.50%	East Midlands
North West Leicestershire	12041	60	0.50%	East Midlands
Mid Suffolk	12066	60	0.50%	East of England
Hillingdon	27406	136	0.50%	London
Mid Devon	9690	48	0.50%	South West
Guildford	17033	84	0.49%	South East
Fareham	15062	74	0.49%	South East
Pendle	9432	46	0.49%	North West
Three Rivers	10472	51	0.49%	East of England
Chiltern	11112	54	0.49%	South East
Sheffield	50829	247	0.49%	Yorkshire and The Humber
South Norfolk	18529	90	0.49%	East of England
Tameside	19403	94	0.48%	North West
Ipswich	16164	78	0.48%	East of England
Nuneaton and Bedworth	13197	63	0.48%	West Midlands
Bournemouth	27352	130	0.48%	South West
Herefordshire, County of	19243	91	0.47%	West Midlands
South Staffordshire	9608	45	0.47%	West Midlands
Craven	7046	33	0.47%	Yorkshire and The Humber
Barrow-in-Furness	8565	40	0.47%	North West
York	25823	120	0.46%	Yorkshire and The Humber
Stoke-on-Trent	23394	108	0.46%	West Midlands
Tendring	20640	95	0.46%	East of England
Southend-on-Sea	23596	108	0.46%	East of England
St	13878	63	0.45%	East of England

LA	House price spatial data	Unable matched	Proportion	Region
Edmundsbury				
Sevenoaks	13242	60	0.45%	South East
South Somerset	20045	90	0.45%	South West
West Somerset	4456	20	0.45%	South West
North Somerset	29243	131	0.45%	South West
Islington	21435	95	0.44%	London
Rutland	4968	22	0.44%	East Midlands
Solihull	23756	105	0.44%	West Midlands
West Lindsey	11369	50	0.44%	East Midlands
Carlisle	12529	55	0.44%	North West
South Gloucestershire	33060	145	0.44%	South West
Wirral	31067	136	0.44%	North West
Hinckley and Bosworth	14426	63	0.44%	East Midlands
Gloucester	16767	73	0.44%	South West
Wolverhampton	18916	82	0.43%	West Midlands
Hart	11373	49	0.43%	South East
Teignbridge	18195	78	0.43%	South West
Spelthorne	12167	52	0.43%	South East
Harborough	11992	51	0.43%	East Midlands
Luton	17788	75	0.42%	East of England
North East Derbyshire	9507	40	0.42%	East Midlands
Epping Forest	15227	64	0.42%	East of England
Wyre	12406	52	0.42%	North West
Woking	14137	59	0.42%	South East
Barking and Dagenham	14707	61	0.41%	London
Breckland	17190	71	0.41%	East of England
Mole Valley	10484	43	0.41%	South East
Basingstoke and Deane	22086	90	0.41%	South East
North East	16490	67	0.41%	Yorkshire and The

LA	House price spatial data	Unable matched	Proportion	Region
Lincolnshire				Humber
Southampton	25664	104	0.41%	South East
Rushmoor	11938	48	0.40%	South East
Chesterfield	10503	42	0.40%	East Midlands
Wycombe	20299	81	0.40%	South East
Cherwell	18598	74	0.40%	South East
Newark and Sherwood	14082	56	0.40%	East Midlands
Kingston upon Hull, City of	22708	90	0.40%	Yorkshire and The Humber
Central Bedfordshire	37229	147	0.39%	East of England
Rossendale	7100	28	0.39%	North West
Sedgemoor	14992	59	0.39%	South West
Copeland	7127	28	0.39%	North West
Dudley	28517	112	0.39%	West Midlands
Epsom and Ewell	10223	40	0.39%	South East
Tunbridge Wells	14570	57	0.39%	South East
Chorley	13569	53	0.39%	North West
Fenland	12332	48	0.39%	East of England
Lancaster	16735	65	0.39%	North West
King's Lynn and West Norfolk	19076	74	0.39%	East of England
Staffordshire Moorlands	10056	39	0.39%	West Midlands
Blackpool	14191	55	0.39%	North West
Gravesham	10051	38	0.38%	South East
Thanet	19202	72	0.37%	South East
Plymouth	29891	112	0.37%	South West
Poole	21101	79	0.37%	South West
Ryedale	6162	23	0.37%	Yorkshire and The Humber
South Kesteven	19089	71	0.37%	East Midlands

LA	House price spatial data	Unable matched	Proportion	Region
East Devon	22059	82	0.37%	South West
Shepway	14560	54	0.37%	South East
Havering	26565	98	0.37%	London
Knowsley	9828	36	0.37%	North West
Cheshire East	46512	170	0.37%	North West
Enfield	26101	95	0.36%	London
Weymouth and Portland	8801	32	0.36%	South West
Greenwich	25652	93	0.36%	London
Arun	24552	89	0.36%	South East
North Kesteven	15175	55	0.36%	East Midlands
Bracknell Forest	15209	55	0.36%	South East
Lewisham	28850	104	0.36%	London
Eastleigh	15819	57	0.36%	South East
Walsall	21465	77	0.36%	West Midlands
Dacorum	18873	67	0.36%	East of England
Swale	17039	60	0.35%	South East
Thurrock	17711	62	0.35%	East of England
Bexley	25417	88	0.35%	London
Ashfield	14272	49	0.34%	East Midlands
Windsor and Maidenhead	17516	60	0.34%	South East
Suffolk Coastal	17005	57	0.34%	East of England
North Norfolk	15284	51	0.33%	East of England
Charnwood	20774	69	0.33%	East Midlands
Kettering	13286	44	0.33%	East Midlands
Mansfield	11524	38	0.33%	East Midlands
Castle Point	10363	34	0.33%	East of England
Hammersmith and Fulham	19945	65	0.33%	London
Newcastle-under-Lyme	12465	40	0.32%	West Midlands
Broadland	15719	50	0.32%	East of England

LA	House price spatial data	Unable matched	Proportion	Region
Dover	14570	46	0.32%	South East
North Lincolnshire	16214	51	0.31%	Yorkshire and The Humber
South Cambridgeshire	19141	60	0.31%	East of England
New Forest	23416	73	0.31%	South East
Kensington and Chelsea	19024	59	0.31%	London
Coventry	31416	97	0.31%	West Midlands
Reigate and Banstead	20222	62	0.31%	South East
Fylde	9810	30	0.31%	North West
Brentwood	10164	31	0.30%	East of England
Havant	14116	43	0.30%	South East
Worcester	12528	38	0.30%	West Midlands
Lewes	13586	41	0.30%	South East
St Albans	19929	60	0.30%	East of England
East Lindsey	16287	49	0.30%	East Midlands
Chelmsford	21628	65	0.30%	East of England
Isle of Wight	20395	61	0.30%	South East
Torridge	9750	29	0.30%	South West
Elmbridge	19702	58	0.29%	South East
High Peak	10226	30	0.29%	East Midlands
East Staffordshire	12294	36	0.29%	West Midlands
East Cambridgeshire	11118	32	0.29%	East of England
Rochford	10107	29	0.29%	East of England
Bromley	41033	117	0.29%	London
Rotherham	23266	66	0.28%	Yorkshire and The Humber
Surrey Heath	10589	30	0.28%	South East
East Riding of Yorkshire	41880	118	0.28%	Yorkshire and The Humber
Redbridge	24242	66	0.27%	London

LA	House price spatial data	Unable matched	Proportion	Region
South Holland	11521	31	0.27%	East Midlands
Croydon	36530	98	0.27%	London
Medway	30620	80	0.26%	South East
Gedling	14232	37	0.26%	East Midlands
Worthing	16935	44	0.26%	South East
Rushcliffe	15102	39	0.26%	East Midlands
Christchurch	7844	20	0.25%	South West
Scarborough	14566	36	0.25%	Yorkshire and The Humber
Waveney	14671	36	0.25%	East of England
Richmond upon Thames	25758	60	0.23%	London
Haringey	19046	43	0.23%	London
Merton	22642	48	0.21%	London
Portsmouth	23684	50	0.21%	South East
South Bucks	7874	16	0.20%	South East
Hastings	12156	24	0.20%	South East
Kingston upon Thames	19247	37	0.19%	London
Broxbourne	11355	21	0.18%	East of England
Brighton and Hove	37863	68	0.18%	South East
Waltham Forest	21839	37	0.17%	London
East Northamptonshire	12275	20	0.16%	East Midlands
Rother	14470	20	0.14%	South East
Tandridge	11019	11	0.10%	South East

Appendix B1

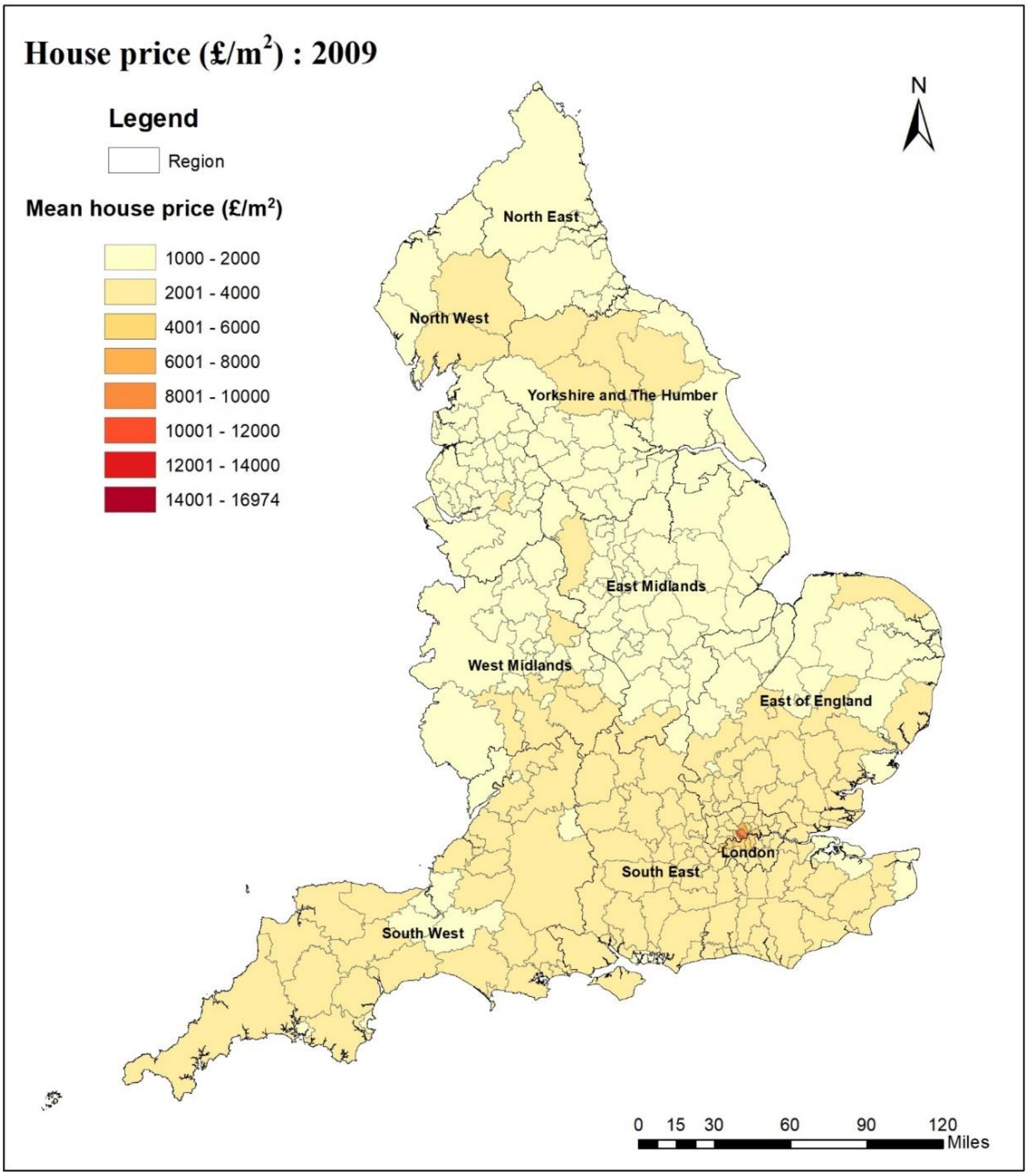


Figure B1.1 The geography of HPM at LA level in 2009

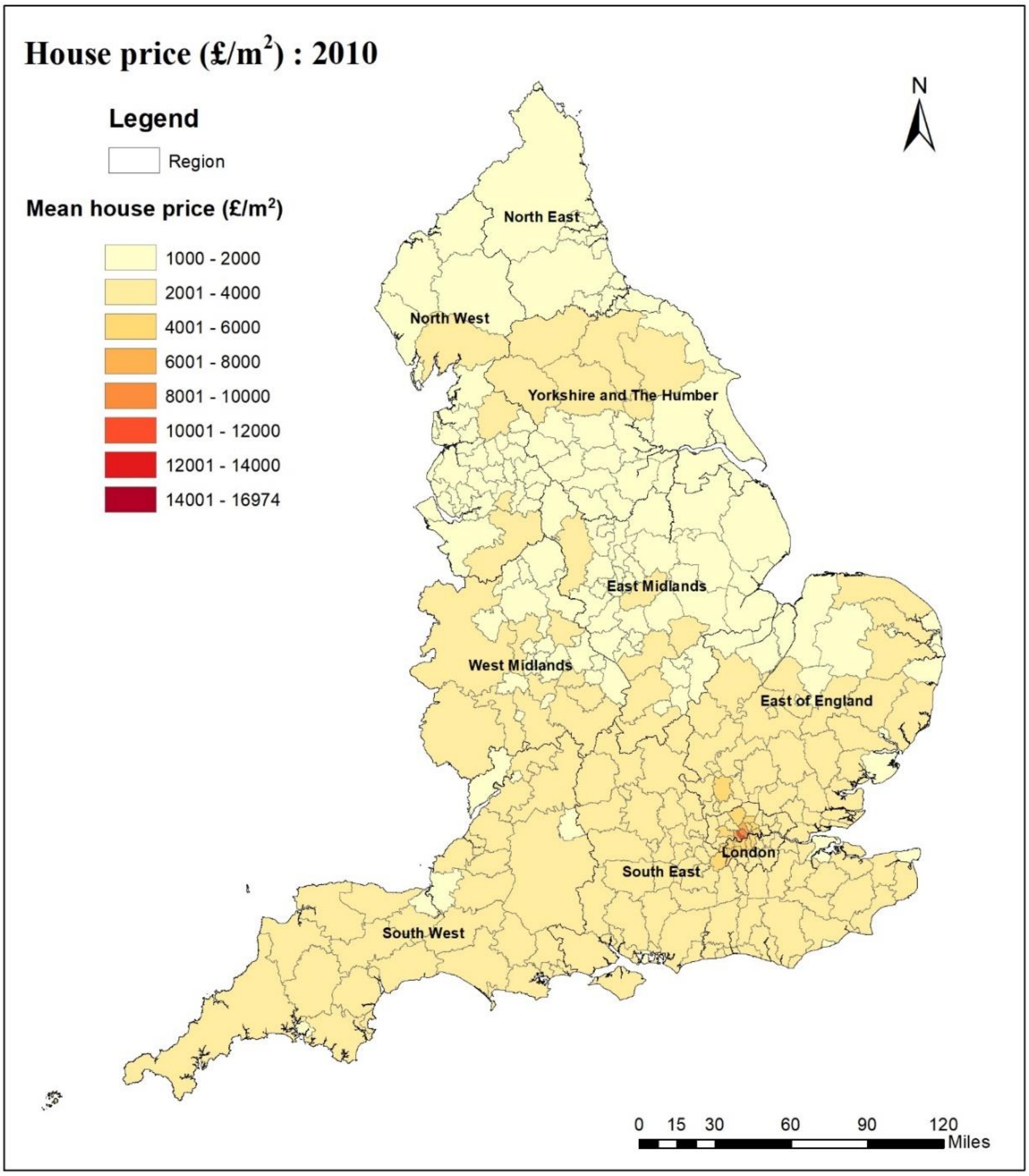


Figure B1.2 The geography of HPM at LA level in 2010

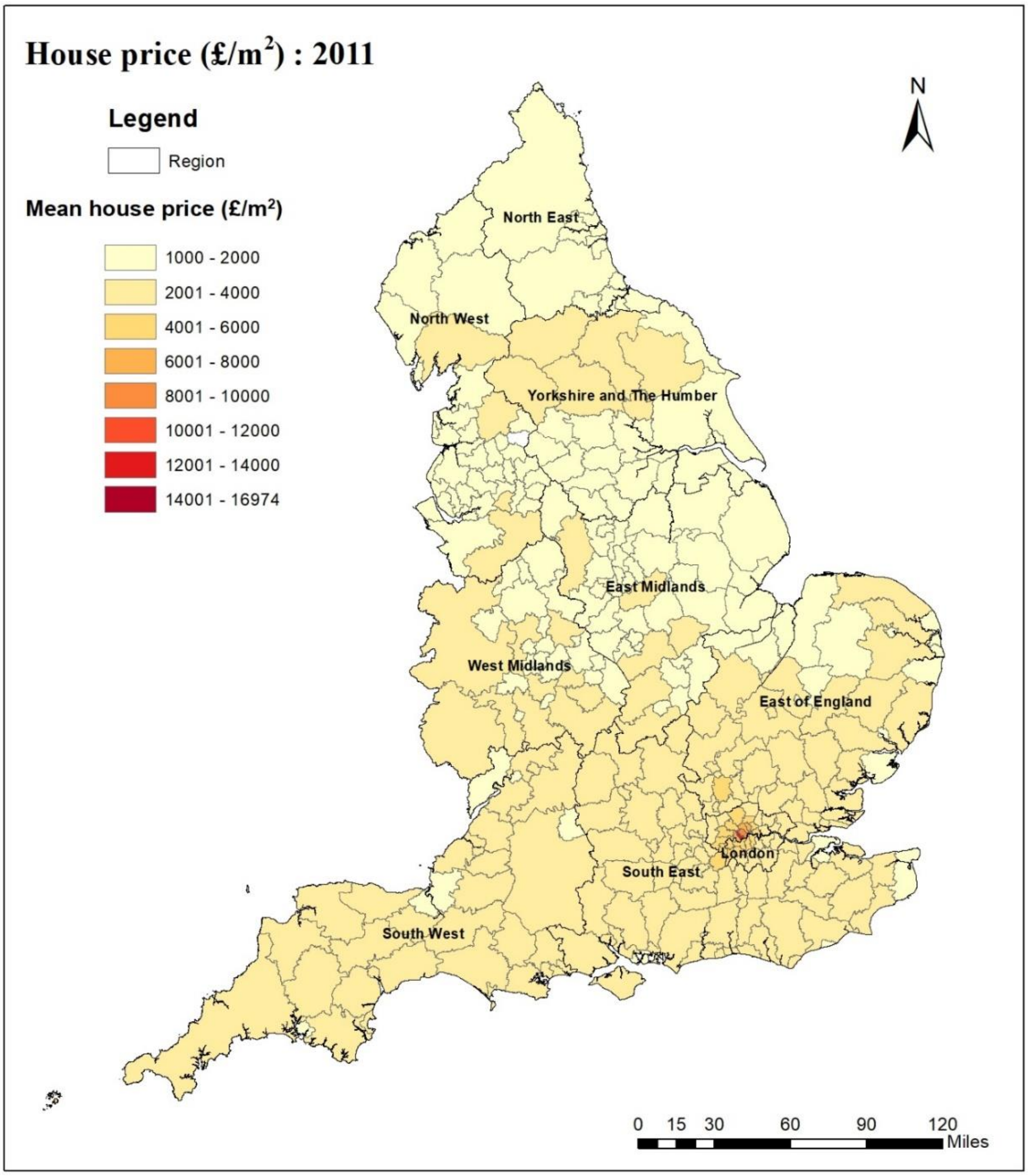


Figure B1.3 The geography of HPM at LA level in 2011

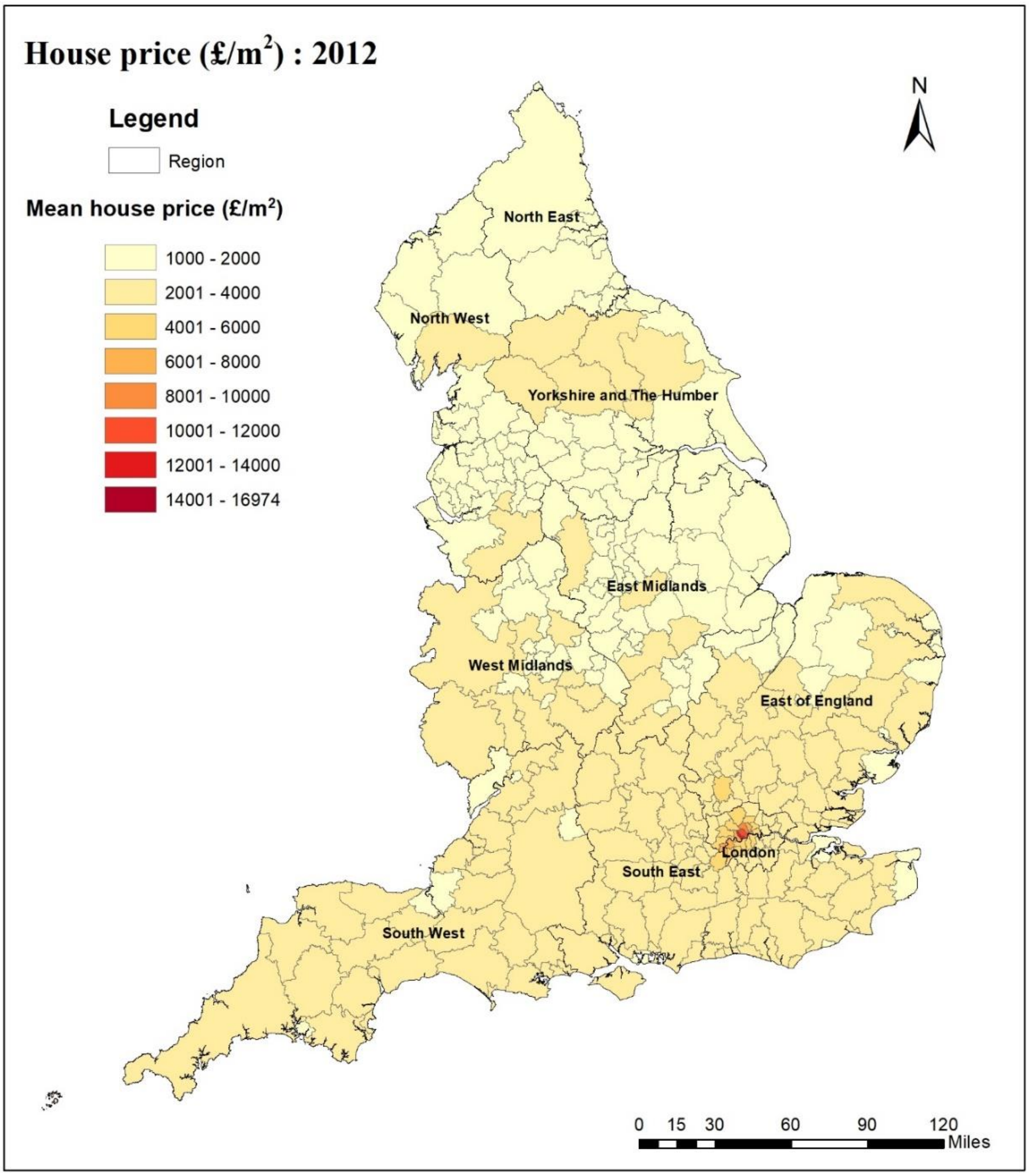


Figure B1.4 The geography of HPM at LA level in 2012

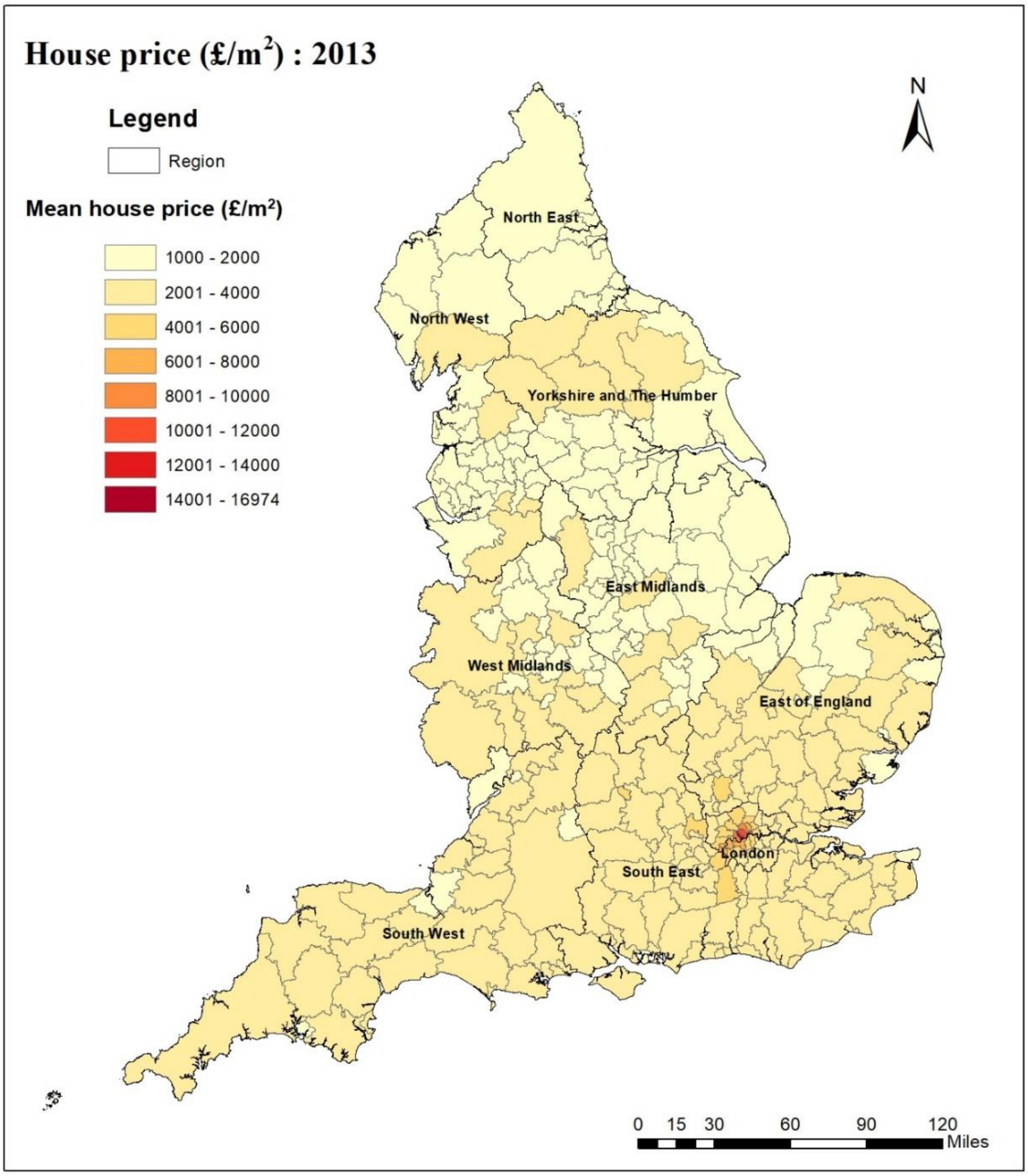


Figure B1.5 The geography of HPM at LA level in 2013

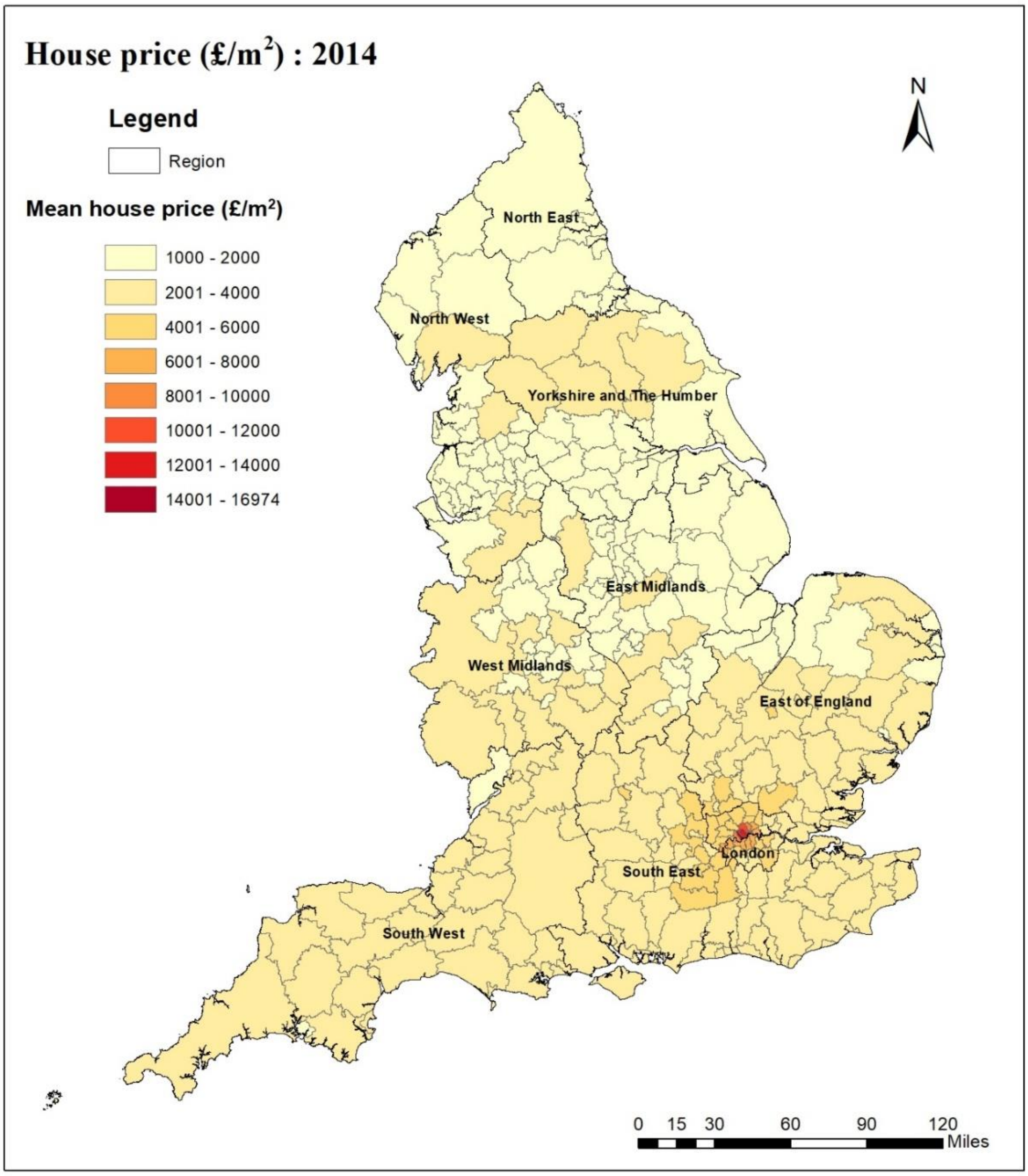


Figure B1.6 The geography of HPM at LA level in 2014

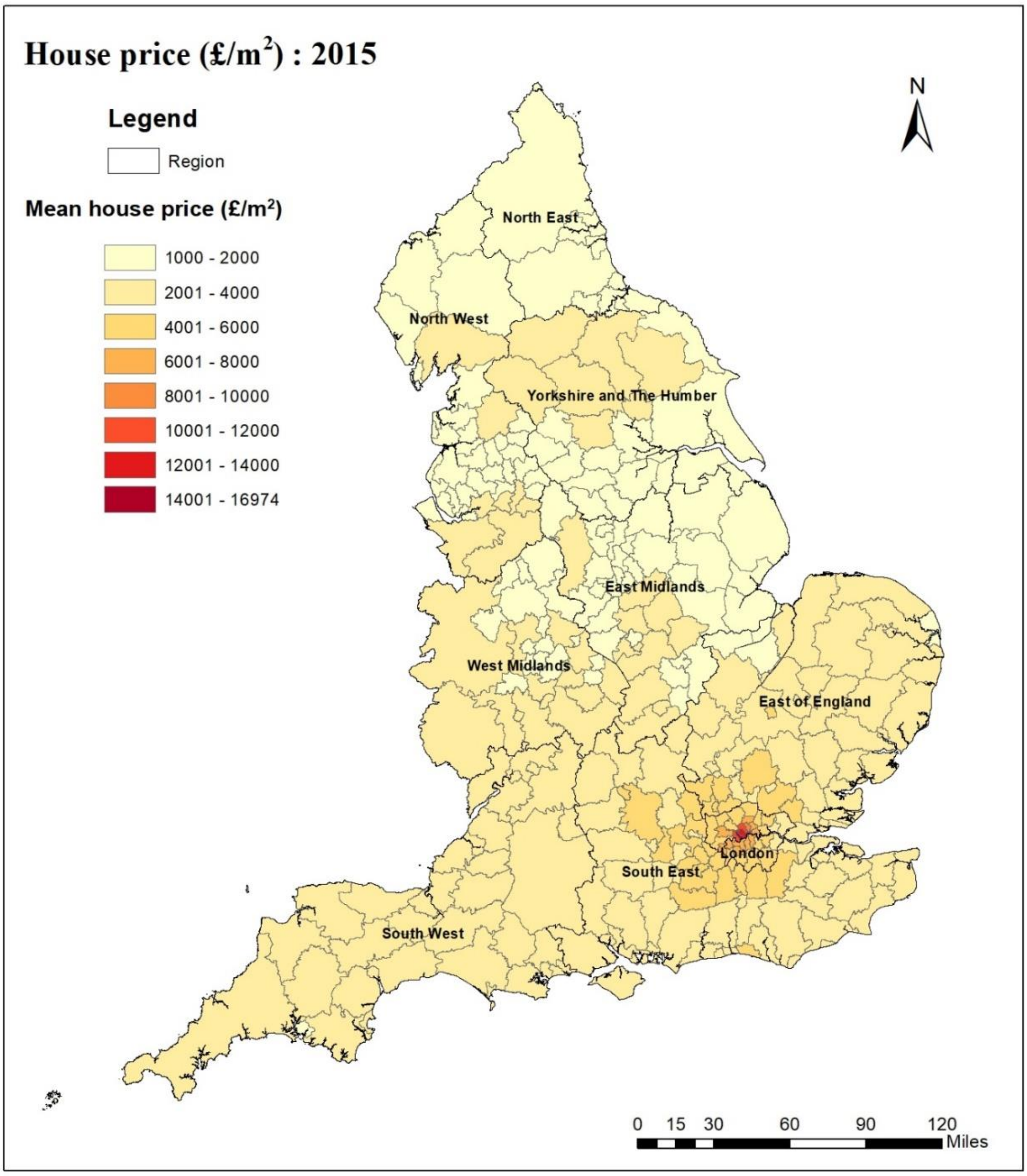


Figure B1.7 The geography of HPM at LA level in 2015

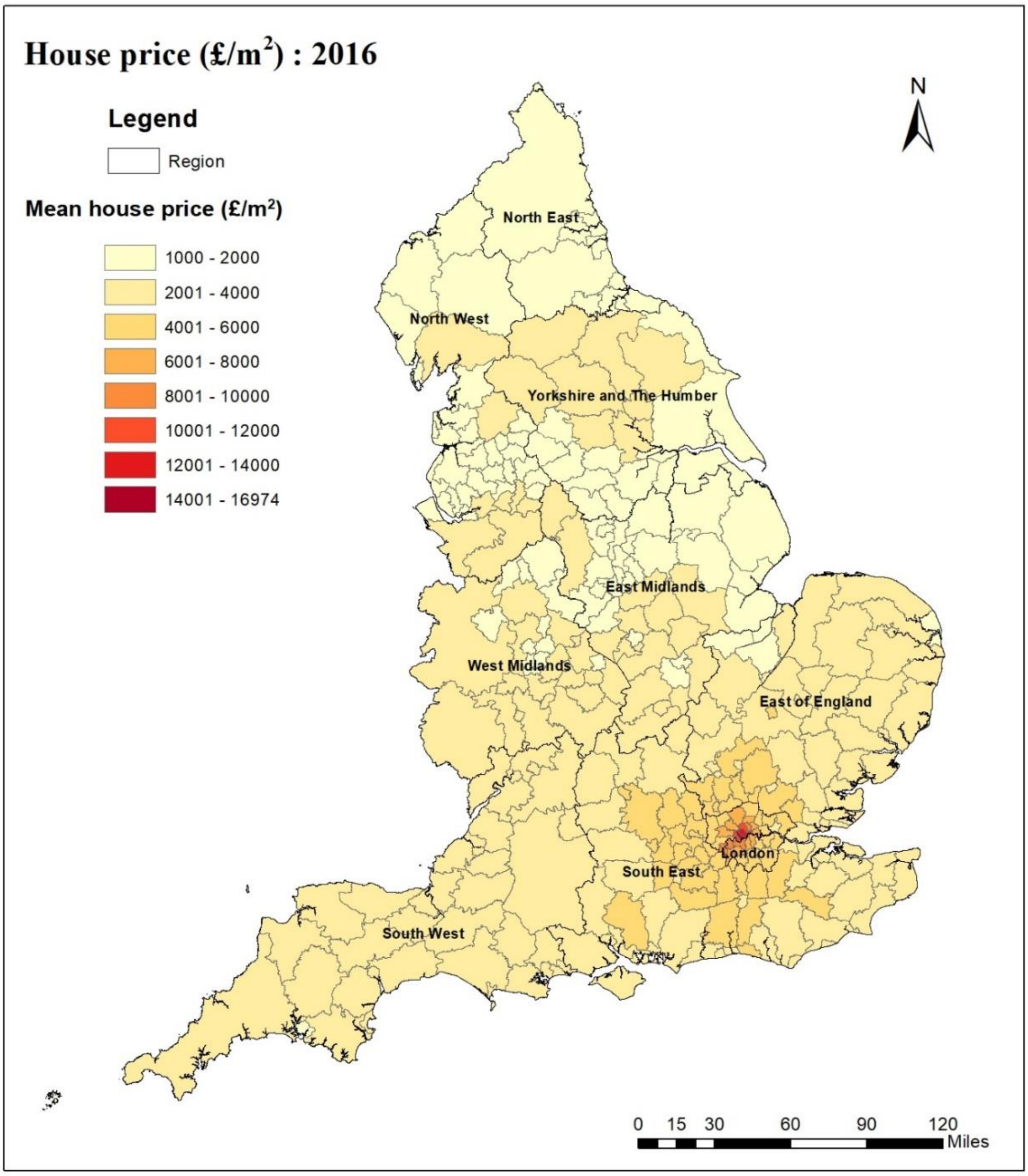


Figure B1.7 The geography of HPM at LA level in 2016

Appendix B2

Table B2 Model result of Models 1 to 3

Model 1			Model 2			Model 3		
Parameter	Estimate	S.E.	Parameter	Estimate	S.E.	Parameter	Estimate	S.E.
β_0 Intercept	7.6991	0.0235	β_0 Intercept	7.6980	0.0235	β_0 Intercept	7.6994	0.0235
σ_l^2 LA level variance	0.1770	0.0141	σ_l^2 LA level variance	0.1768	0.0141	σ_l^2 LA level variance	0.1771	0.0141
σ_m^2 MSOA level variance	0.0364	0.0007	σ_m^2 MSOA level variance	0.0361	0.0007	σ_m^2 MSOA level variance	0.0353	0.0007
σ_q^2 Quarter level variance	0.0140	0.0001	σ_{hy}^2 Half-year level variance	0.0142	0.0001	σ_y^2 Year level variance	0.0143	0.0001
σ_e^2 Individual level variance	0.0735	0.0000	σ_e^2 Individual level variance	0.0737	0.0000	σ_e^2 Individual level variance	0.0743	0.0000
Deviance	1,428,443		Deviance	1,338,665		Deviance	1,287,883	

Appendix B3

Table B3 Five-group clusters result on LA house price trend

Group	LA name	LA code	Region	HPM percentage increase	Starting-price
5	City of London	E09000001	London	9.63%	5300.34
	Camden	E09000007	London	8.59%	5904.71
	Hammersmith and Fulham	E09000013	London	8.70%	5389.90
	Islington	E09000019	London	9.08%	4717.28
	Kensington and Chelsea	E09000020	London	8.48%	8503.86
	Westminster	E09000033	London	9.08%	7118.36
4	Slough	E06000039	South East	7.11%	2335.20
	Windsor and Maidenhead	E06000040	South East	5.62%	3168.27
	Chiltern	E07000005	South East	5.04%	3270.22
	South Bucks	E07000006	South East	5.66%	3172.86
	Cambridge	E07000008	East of England	7.37%	2984.02
	Epping Forest	E07000072	East of England	5.91%	2911.85
	Dacorum	E07000096	East of England	6.46%	2636.96
	Hertsmere	E07000098	East of England	6.77%	3022.75
	Three Rivers	E07000102	East of England	6.61%	3106.86
	Watford	E07000103	East of England	7.54%	2639.21
	Oxford	E07000178	South East	6.17%	3096.90
	Elmbridge	E07000207	South East	5.85%	3656.36

	Epsom and Ewell	E07000208	South East	6.75%	3051.84
	Guildford	E07000209	South East	5.44%	3138.13
	Mole Valley	E07000210	South East	5.33%	3260.50
	Reigate and Banstead	E07000211	South East	6.10%	2771.25
	Runnymede	E07000212	South East	5.60%	3041.92
	Spelthorne	E07000213	South East	6.28%	2809.32
	Woking	E07000217	South East	5.62%	3001.58
	St Albans	E07000240	East of England	6.33%	3410.82
	Barking and Dagenham	E09000002	London	8.00%	2009.57
	Barnet	E09000003	London	7.07%	3575.91
	Bexley	E09000004	London	7.49%	2247.84
	Brent	E09000005	London	8.35%	3172.50
	Bromley	E09000006	London	7.59%	2757.34
	Croydon	E09000008	London	8.09%	2357.04
	Ealing	E09000009	London	8.56%	3217.83
	Enfield	E09000010	London	7.60%	2704.12
	Greenwich	E09000011	London	8.78%	2418.59
	Hackney	E09000012	London	10.53%	3478.93
	Haringey	E09000014	London	9.48%	3163.14
	Harrow	E09000015	London	7.14%	3102.59
	Hillingdon	E09000017	London	7.20%	2883.58
	Hounslow	E09000018	London	7.17%	3014.48

	Kingston upon Thames	E09000021	London	7.60%	3415.49
	Lambeth	E09000022	London	10.33%	3511.14
	Lewisham	E09000023	London	10.33%	2611.76
	Merton	E09000024	London	8.40%	3330.67
	Newham	E09000025	London	9.01%	2225.54
	Redbridge	E09000026	London	6.63%	2663.98
	Richmond upon Thames	E09000027	London	7.40%	4570.20
	Southwark	E09000028	London	9.71%	3452.23
	Sutton	E09000029	London	7.54%	2681.95
	Tower Hamlets	E09000030	London	8.15%	3754.78
	Waltham Forest	E09000031	London	10.81%	2484.04
	Wandsworth	E09000032	London	8.52%	4312.02
3	Bath and North East Somerset	E06000022	South West	4.59%	2280.09
	Bristol, City of	E06000023	South West	6.70%	1772.51
	South Gloucestershire	E06000025	South West	4.99%	1975.75
	Bournemouth	E06000028	South West	3.89%	2165.61
	Poole	E06000029	South West	3.66%	2330.84
	Luton	E06000032	East of England	6.17%	1703.23
	Southend-on-Sea	E06000033	East of England	5.45%	1984.10
	Thurrock	E06000034	East of England	6.08%	2003.92
	Medway	E06000035	South East	6.14%	1685.22
	Bracknell Forest	E06000036	South East	6.27%	2483.91

	West Berkshire	E06000037	South East	4.71%	2585.42
	Reading	E06000038	South East	6.34%	2311.10
	Wokingham	E06000041	South East	5.46%	2812.99
	Milton Keynes	E06000042	South East	6.09%	1759.47
	Brighton and Hove	E06000043	South East	5.74%	2803.30
	Bedford	E06000055	East of England	5.11%	1831.11
	Central Bedfordshire	E06000056	East of England	5.71%	2020.60
	Aylesbury Vale	E07000004	South East	4.82%	2406.24
	Wycombe	E07000007	South East	5.21%	2786.93
	East Cambridgeshire	E07000009	East of England	5.49%	1881.02
	Huntingdonshire	E07000011	East of England	4.84%	1801.29
	South Cambridgeshire	E07000012	East of England	5.34%	2385.11
	Christchurch	E07000048	South West	3.86%	2542.84
	East Dorset	E07000049	South West	3.29%	2515.23
	Lewes	E07000063	South East	4.34%	2484.72
	Wealden	E07000065	South East	3.76%	2509.87
	Basildon	E07000066	East of England	5.69%	2040.63
	Braintree	E07000067	East of England	4.77%	2044.60
	Brentwood	E07000068	East of England	5.03%	2917.12
	Castle Point	E07000069	East of England	4.59%	2180.53
	Chelmsford	E07000070	East of England	5.00%	2458.58
	Colchester	E07000071	East of England	4.76%	1963.25

	Harlow	E07000073	East of England	6.55%	1903.16
	Maldon	E07000074	East of England	4.45%	2173.47
	Rochford	E07000075	East of England	4.55%	2282.34
	Uttlesford	E07000077	East of England	3.72%	2646.00
	Cheltenham	E07000078	South West	3.94%	2116.59
	Basingstoke and Deane	E07000084	South East	5.24%	2280.65
	East Hampshire	E07000085	South East	4.28%	2556.20
	Eastleigh	E07000086	South East	4.26%	2262.27
	Fareham	E07000087	South East	3.89%	2237.96
	Hart	E07000089	South East	4.98%	2802.82
	New Forest	E07000091	South East	3.59%	2561.36
	Rushmoor	E07000092	South East	5.90%	2217.11
	Test Valley	E07000093	South East	3.58%	2380.76
	Winchester	E07000094	South East	4.28%	2826.24
	Broxbourne	E07000095	East of England	6.05%	2501.08
	North Hertfordshire	E07000099	East of England	5.71%	2416.19
	Ashford	E07000105	South East	4.63%	2091.42
	Canterbury	E07000106	South East	4.92%	2183.59
	Dartford	E07000107	South East	7.22%	2121.76
	Gravesham	E07000109	South East	6.10%	1993.13
	Maidstone	E07000110	South East	4.98%	2199.27
	Sevenoaks	E07000111	South East	5.27%	2744.19

	Swale	E07000113	South East	5.27%	1700.93
	Tonbridge and Malling	E07000115	South East	5.17%	2459.36
	Tunbridge Wells	E07000116	South East	5.00%	2657.97
	South Northamptonshire	E07000155	East Midlands	4.13%	2081.49
	Cherwell	E07000177	South East	4.82%	2307.68
	South Oxfordshire	E07000179	South East	4.46%	3000.07
	Vale of White Horse	E07000180	South East	4.10%	2637.05
	West Oxfordshire	E07000181	South East	3.79%	2657.46
	St Edmundsbury	E07000204	East of England	4.87%	1918.58
	Surrey Heath	E07000214	South East	5.52%	2712.74
	Tandridge	E07000215	South East	5.25%	2808.31
	Waverley	E07000216	South East	4.38%	3228.84
	Warwick	E07000222	West Midlands	5.07%	2182.40
	Adur	E07000223	South East	5.43%	2380.25
	Arun	E07000224	South East	3.99%	2238.59
	Chichester	E07000225	South East	3.56%	2842.30
	Crawley	E07000226	South East	6.61%	2120.46
	Horsham	E07000227	South East	4.32%	2779.56
	Mid Sussex	E07000228	South East	4.91%	2696.49
	Worthing	E07000229	South East	5.03%	2190.12
	Welwyn Hatfield	E07000241	East of England	6.10%	2622.18
	East Hertfordshire	E07000242	East of England	5.24%	2824.06

	Stevenage	E07000243	East of England	6.55%	1882.71
	Havering	E09000016	London	6.52%	2381.45
2	Warrington	E06000007	North West	2.62%	1547.51
	York	E06000014	Yorkshire and The Humber	3.51%	2100.08
	Derby	E06000015	East Midlands	3.19%	1340.81
	Leicester	E06000016	East Midlands	3.64%	1312.37
	Rutland	E06000017	East Midlands	3.13%	1899.28
	Nottingham	E06000018	East Midlands	3.53%	1147.66
	Herefordshire, County of	E06000019	West Midlands	2.02%	1890.13
	North Somerset	E06000024	South West	4.05%	1940.58
	Plymouth	E06000026	South West	2.37%	1579.89
	Torbay	E06000027	South West	1.91%	1802.19
	Swindon	E06000030	South West	4.83%	1593.36
	Peterborough	E06000031	East of England	3.61%	1393.01
	Portsmouth	E06000044	South East	4.28%	1672.56
	Southampton	E06000045	South East	3.72%	1876.91
	Cheshire East	E06000049	North West	2.59%	1757.19
	Cheshire West and Chester	E06000050	North West	2.08%	1664.53
	Cornwall	E06000052	South West	1.52%	2095.88
	Isles of Scilly	E06000053	South West	1.50%	2503.25
	Wiltshire	E06000054	South West	3.43%	2099.67
	Fenland	E07000010	East of England	3.70%	1377.56

	Amber Valley	E07000032	East Midlands	3.00%	1397.42
	Chesterfield	E07000034	East Midlands	2.99%	1354.72
	Derbyshire Dales	E07000035	East Midlands	2.40%	2060.93
	Erewash	E07000036	East Midlands	3.95%	1312.91
	High Peak	E07000037	East Midlands	2.18%	1614.43
	North East Derbyshire	E07000038	East Midlands	2.24%	1549.74
	South Derbyshire	E07000039	East Midlands	3.02%	1553.85
	East Devon	E07000040	South West	2.11%	2338.05
	Exeter	E07000041	South West	3.01%	2177.75
	Mid Devon	E07000042	South West	1.74%	1975.87
	North Devon	E07000043	South West	1.95%	1996.27
	South Hams	E07000044	South West	2.07%	2519.83
	Teignbridge	E07000045	South West	2.02%	2140.63
	West Devon	E07000047	South West	1.60%	1989.39
	North Dorset	E07000050	South West	2.17%	2163.27
	Purbeck	E07000051	South West	2.83%	2460.83
	West Dorset	E07000052	South West	2.22%	2356.93
	Weymouth and Portland	E07000053	South West	2.13%	1997.51
	Eastbourne	E07000061	South East	4.02%	1970.89
	Hastings	E07000062	South East	4.14%	1665.97
	Rother	E07000064	South East	3.08%	2217.77
	Tendring	E07000076	East of England	3.42%	1709.15

	Cotswold	E07000079	South West	2.49%	2635.90
	Forest of Dean	E07000080	South West	2.10%	1758.26
	Gloucester	E07000081	South West	4.02%	1586.49
	Stroud	E07000082	South West	2.78%	2143.15
	Tewkesbury	E07000083	South West	3.37%	2081.45
	Gosport	E07000088	South East	3.62%	1712.79
	Havant	E07000090	South East	3.90%	1959.43
	Dover	E07000108	South East	4.65%	1720.90
	Shepway	E07000112	South East	3.43%	1870.29
	Thanet	E07000114	South East	4.45%	1651.01
	Blaby	E07000129	East Midlands	3.50%	1639.80
	Charnwood	E07000130	East Midlands	2.91%	1683.92
	Harborough	E07000131	East Midlands	3.96%	1818.85
	Hinckley and Bosworth	E07000132	East Midlands	2.90%	1642.87
	Melton	E07000133	East Midlands	2.93%	1677.60
	North West Leicestershire	E07000134	East Midlands	2.91%	1527.66
	Oadby and Wigston	E07000135	East Midlands	4.00%	1613.85
	Boston	E07000136	East Midlands	2.69%	1261.24
	Lincoln	E07000138	East Midlands	3.38%	1325.06
	North Kesteven	E07000139	East Midlands	3.05%	1494.61
	South Holland	E07000140	East Midlands	3.80%	1334.71
	South Kesteven	E07000141	East Midlands	3.65%	1501.11

	Breckland	E07000143	East of England	3.69%	1620.58
	Broadland	E07000144	East of England	3.78%	1856.37
	Great Yarmouth	E07000145	East of England	2.61%	1414.94
	King's Lynn and West Norfolk	E07000146	East of England	3.07%	1591.24
	North Norfolk	E07000147	East of England	3.16%	1881.36
	Norwich	E07000148	East of England	4.79%	1676.93
	South Norfolk	E07000149	East of England	3.58%	1864.30
	Corby	E07000150	East Midlands	5.10%	1193.08
	Daventry	E07000151	East Midlands	3.49%	1818.29
	East Northamptonshire	E07000152	East Midlands	4.39%	1553.82
	Kettering	E07000153	East Midlands	4.38%	1404.27
	Northampton	E07000154	East Midlands	4.76%	1494.25
	Wellingborough	E07000156	East Midlands	4.57%	1383.18
	Harrogate	E07000165	Yorkshire and The Humber	2.73%	2216.29
	Selby	E07000169	Yorkshire and The Humber	2.23%	1648.10
	Ashfield	E07000170	East Midlands	3.40%	1108.12
	Broxtowe	E07000172	East Midlands	3.66%	1427.64
	Gedling	E07000173	East Midlands	3.19%	1434.91
	Mansfield	E07000174	East Midlands	2.87%	1083.27
	Newark and Sherwood	E07000175	East Midlands	2.70%	1462.98
	Rushcliffe	E07000176	East Midlands	3.85%	1857.15
	Mendip	E07000187	South West	3.27%	2015.75

	Sedgemoor	E07000188	South West	2.19%	1765.91
	South Somerset	E07000189	South West	1.83%	1915.95
	Taunton Deane	E07000190	South West	1.98%	1907.10
	Cannock Chase	E07000192	West Midlands	2.89%	1437.15
	East Staffordshire	E07000193	West Midlands	2.68%	1477.84
	Lichfield	E07000194	West Midlands	2.84%	1864.39
	Newcastle-under-Lyme	E07000195	West Midlands	2.52%	1286.20
	South Staffordshire	E07000196	West Midlands	1.81%	1856.91
	Stafford	E07000197	West Midlands	2.09%	1655.64
	Staffordshire Moorlands	E07000198	West Midlands	2.20%	1542.86
	Tamworth	E07000199	West Midlands	3.60%	1509.99
	Babergh	E07000200	East of England	3.94%	2018.71
	Forest Heath	E07000201	East of England	4.26%	1679.72
	Ipswich	E07000202	East of England	4.93%	1506.79
	Mid Suffolk	E07000203	East of England	3.68%	1902.95
	Suffolk Coastal	E07000205	East of England	3.51%	2042.19
	Waveney	E07000206	East of England	3.06%	1600.40
	North Warwickshire	E07000218	West Midlands	2.34%	1661.17
	Nuneaton and Bedworth	E07000219	West Midlands	2.88%	1420.58
	Rugby	E07000220	West Midlands	4.56%	1590.30
	Stratford-on-Avon	E07000221	West Midlands	3.09%	2323.84
	Bromsgrove	E07000234	West Midlands	2.69%	2020.86

	Malvern Hills	E07000235	West Midlands	2.34%	1989.08
	Redditch	E07000236	West Midlands	3.41%	1586.97
	Worcester	E07000237	West Midlands	2.64%	1771.12
	Wychavon	E07000238	West Midlands	1.92%	2060.15
	Bury	E08000002	North West	2.35%	1389.24
	Manchester	E08000003	North West	3.54%	1333.56
	Salford	E08000006	North West	2.90%	1215.35
	Stockport	E08000007	North West	3.66%	1706.13
	Trafford	E08000009	North West	4.25%	1825.37
	Sheffield	E08000019	Yorkshire and The Humber	2.42%	1387.96
	Birmingham	E08000025	West Midlands	2.99%	1411.21
	Coventry	E08000026	West Midlands	4.12%	1365.46
	Solihull	E08000029	West Midlands	3.83%	1848.33
1	Hartlepool	E06000001	North East	-1.05%	1070.50
	Middlesbrough	E06000002	North East	-0.22%	1158.16
	Redcar and Cleveland	E06000003	North East	-0.09%	1252.60
	Stockton-on-Tees	E06000004	North East	0.82%	1254.81
	Darlington	E06000005	North East	0.32%	1282.89
	Halton	E06000006	North West	1.59%	1234.88
	Blackburn with Darwen	E06000008	North West	-0.08%	1098.65
	Blackpool	E06000009	North West	-0.22%	1110.05
	Kingston upon Hull, City of	E06000010	Yorkshire and The Humber	1.78%	996.15

	East Riding of Yorkshire	E06000011	Yorkshire and The Humber	1.20%	1526.97
	North East Lincolnshire	E06000012	Yorkshire and The Humber	0.82%	1078.86
	North Lincolnshire	E06000013	Yorkshire and The Humber	1.28%	1176.28
	Telford and Wrekin	E06000020	West Midlands	1.65%	1442.94
	Stoke-on-Trent	E06000021	West Midlands	2.45%	1039.25
	Isle of Wight	E06000046	South East	1.47%	2003.62
	County Durham	E06000047	North East	-0.52%	1102.73
	Shropshire	E06000051	West Midlands	1.35%	1881.00
	Northumberland	E06000057	North East	0.34%	1445.22
	Allerdale	E07000026	North West	-0.56%	1421.08
	Barrow-in-Furness	E07000027	North West	0.54%	1219.27
	Carlisle	E07000028	North West	0.00%	1368.54
	Copeland	E07000029	North West	0.79%	1148.91
	Eden	E07000030	North West	-0.58%	1816.43
	South Lakeland	E07000031	North West	0.39%	2287.14
	Bolsover	E07000033	East Midlands	2.50%	1069.47
	Torridge	E07000046	South West	1.50%	1923.77
	Burnley	E07000117	North West	-0.61%	991.28
	Chorley	E07000118	North West	0.82%	1569.32
	Fylde	E07000119	North West	1.01%	1705.19
	Hyndburn	E07000120	North West	0.26%	1021.27
	Lancaster	E07000121	North West	0.64%	1582.51

	Pendle	E07000122	North West	0.46%	1004.36
	Preston	E07000123	North West	0.41%	1365.61
	Ribble Valley	E07000124	North West	0.14%	1889.64
	Rossendale	E07000125	North West	0.53%	1240.45
	South Ribble	E07000126	North West	0.83%	1608.47
	West Lancashire	E07000127	North West	1.13%	1485.67
	Wyre	E07000128	North West	0.02%	1587.97
	East Lindsey	E07000137	East Midlands	1.70%	1430.82
	West Lindsey	E07000142	East Midlands	2.03%	1331.36
	Craven	E07000163	Yorkshire and The Humber	1.24%	1945.37
	Hambleton	E07000164	Yorkshire and The Humber	1.35%	2026.37
	Richmondshire	E07000166	Yorkshire and The Humber	0.91%	1849.39
	Ryedale	E07000167	Yorkshire and The Humber	0.45%	1978.35
	Scarborough	E07000168	Yorkshire and The Humber	-0.30%	1693.55
	Bassetlaw	E07000171	East Midlands	2.12%	1218.47
	West Somerset	E07000191	South West	1.35%	1973.46
	Wyre Forest	E07000239	West Midlands	1.59%	1658.12
	Bolton	E08000001	North West	1.43%	1213.67
	Oldham	E08000004	North West	1.09%	1263.27
	Rochdale	E08000005	North West	0.69%	1263.09
	Tameside	E08000008	North West	2.24%	1288.25
	Wigan	E08000010	North West	1.43%	1222.24

	Knowsley	E08000011	North West	1.01%	1154.43
	Liverpool	E08000012	North West	1.22%	1139.63
	St. Helens	E08000013	North West	0.63%	1267.85
	Sefton	E08000014	North West	1.14%	1348.58
	Wirral	E08000015	North West	1.43%	1353.64
	Barnsley	E08000016	Yorkshire and The Humber	0.91%	1188.71
	Doncaster	E08000017	Yorkshire and The Humber	1.33%	1172.79
	Rotherham	E08000018	Yorkshire and The Humber	1.94%	1185.19
	Newcastle upon Tyne	E08000021	North East	1.20%	1424.48
	North Tyneside	E08000022	North East	1.79%	1451.63
	South Tyneside	E08000023	North East	0.67%	1255.26
	Sunderland	E08000024	North East	-0.26%	1211.06
	Dudley	E08000027	West Midlands	2.02%	1502.24
	Sandwell	E08000028	West Midlands	2.33%	1260.95
	Walsall	E08000030	West Midlands	2.27%	1322.01
	Wolverhampton	E08000031	West Midlands	2.15%	1229.16
	Bradford	E08000032	Yorkshire and The Humber	-0.18%	1321.02
	Calderdale	E08000033	Yorkshire and The Humber	0.52%	1362.14
	Kirklees	E08000034	Yorkshire and The Humber	1.23%	1410.51
	Leeds	E08000035	Yorkshire and The Humber	1.91%	1577.03
	Wakefield	E08000036	Yorkshire and The Humber	1.14%	1357.93
	Gateshead	E08000037	North East	0.97%	1335.78

Appendix C1

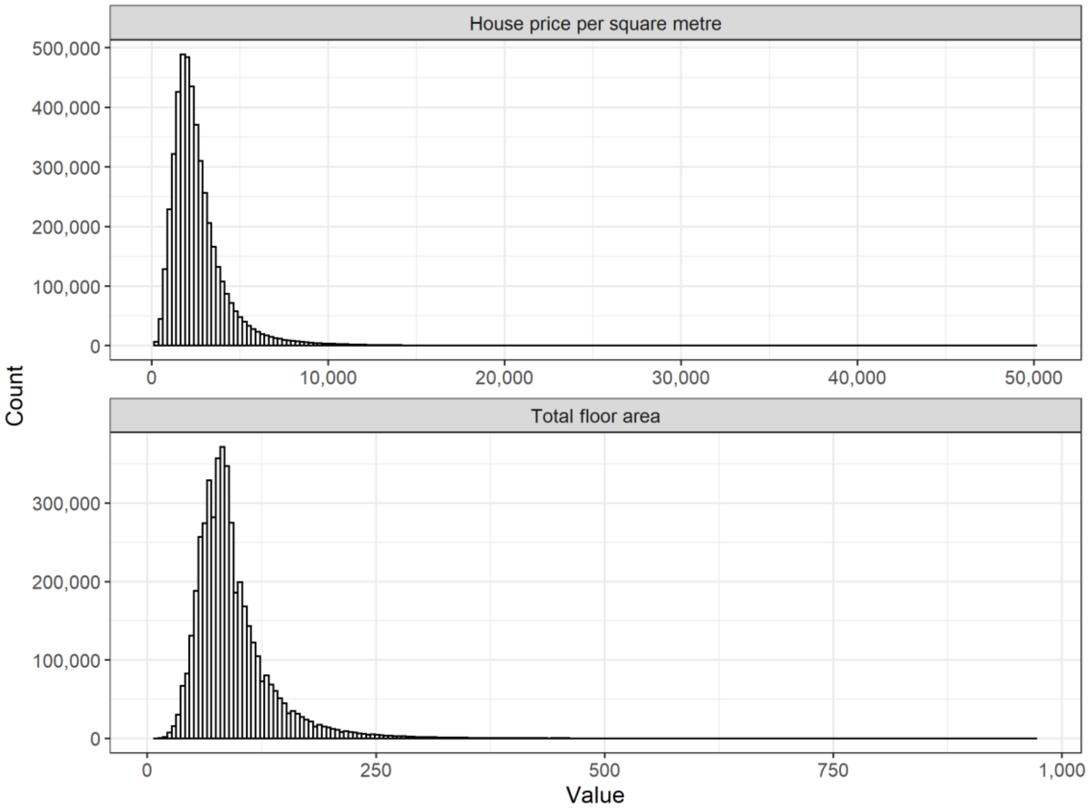


Figure C1 The distribution of HPM and total floor area in England between 2009 and 2016

Table C1.1 A summary of the most common sold property size by property type in England, 2009-2016

Property type	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**
Detached	96	104	87	100	109	100	110	95
Semi-Detached	84	84	84	84	84	84	84	84
Terraced	78	78	78	78	80	80	80	78
Flats/Maisonettes	60	60	60	60	60	60	60	60

Table C1.2 LAs which have lower than 30 annual sample size

Property type	**LA**	**Region**
	Isles of Scilly	South West
	City of London	
	Barking and Dagenham	
	Camden	
	Hackney	
	Hammersmith and Fulham	
	Haringey	
	Islington	
	Kensington and Chelsea	
	Lambeth	
Detached	Newham	London

Property type	**LA**	**Region**
	Southwark	
	Tower Hamlets	
	Waltham Forest	
	Westminster	
	Isles of Scilly	South West
	Hackney	
	Islington	
	Kensington and Chelsea	
	Tower Hamlets	
Semi-detached	Westminster	London
	Isles of Scilly	South West
Terraced	City of London	London
	Isles of Scilly	South West
	Rossendale	
	Barrow-in-Furness	
	Copeland	
	Burnley	
	Hyndburn	
	Pendle	
	Ribble Valle	North West
	North Lincolnshire	
	Ryedale	
Flats/maisonettes	Selby	Yorkshire and The Humber

Property type	**LA**	**Region**
	Richmondshire	
	Oadby and Wigston	
	Bassetlaw	
	South Derbyshire	
	Ashfield	
	Rutland	
	North Kesteven	
	South Holland	
	Melton	
	Blaby	
	North East Derbyshire	
	North West Leicestershire	
	West Lindsey	
	Boston	
	Bolsover	East Midlands
	Staffordshire Moorlands	West Midlands

Appendix C2

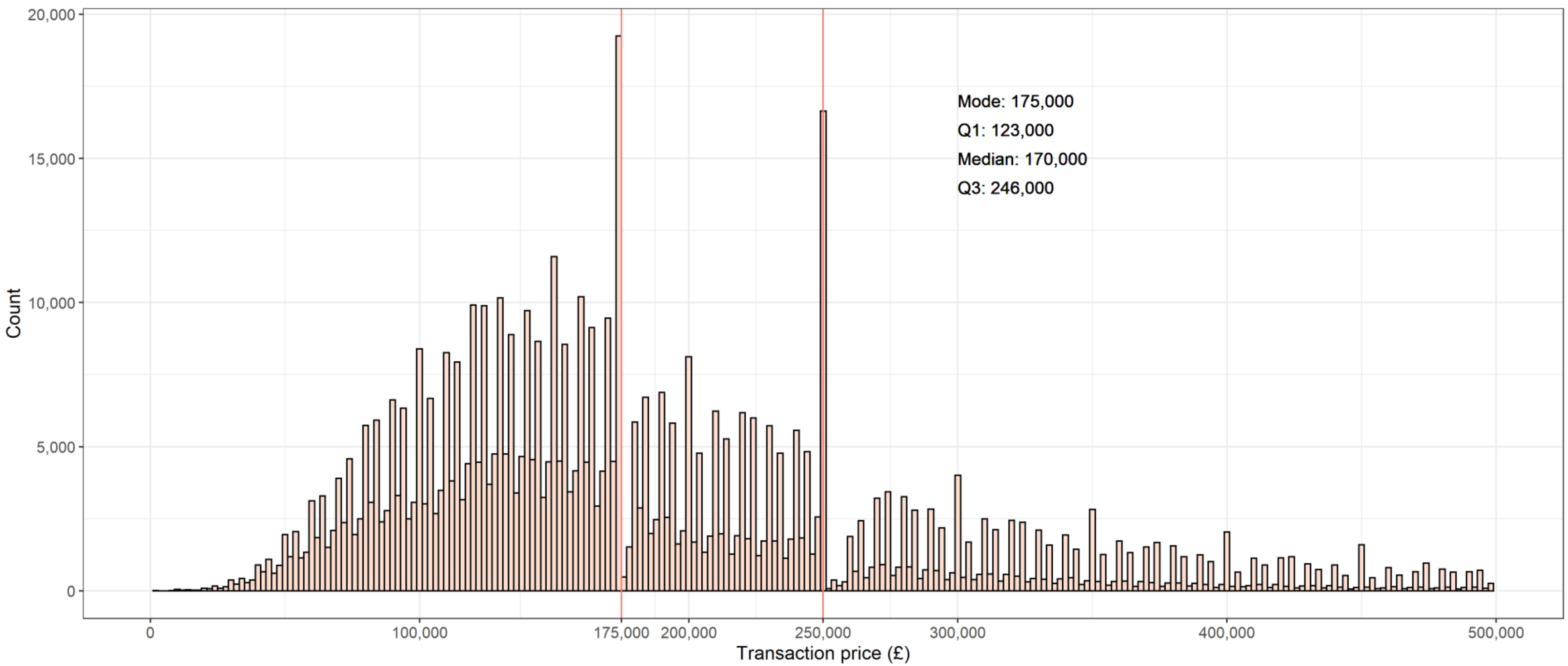


Figure C2.1 The histogram of 2009 TPs in England for value below £500,000

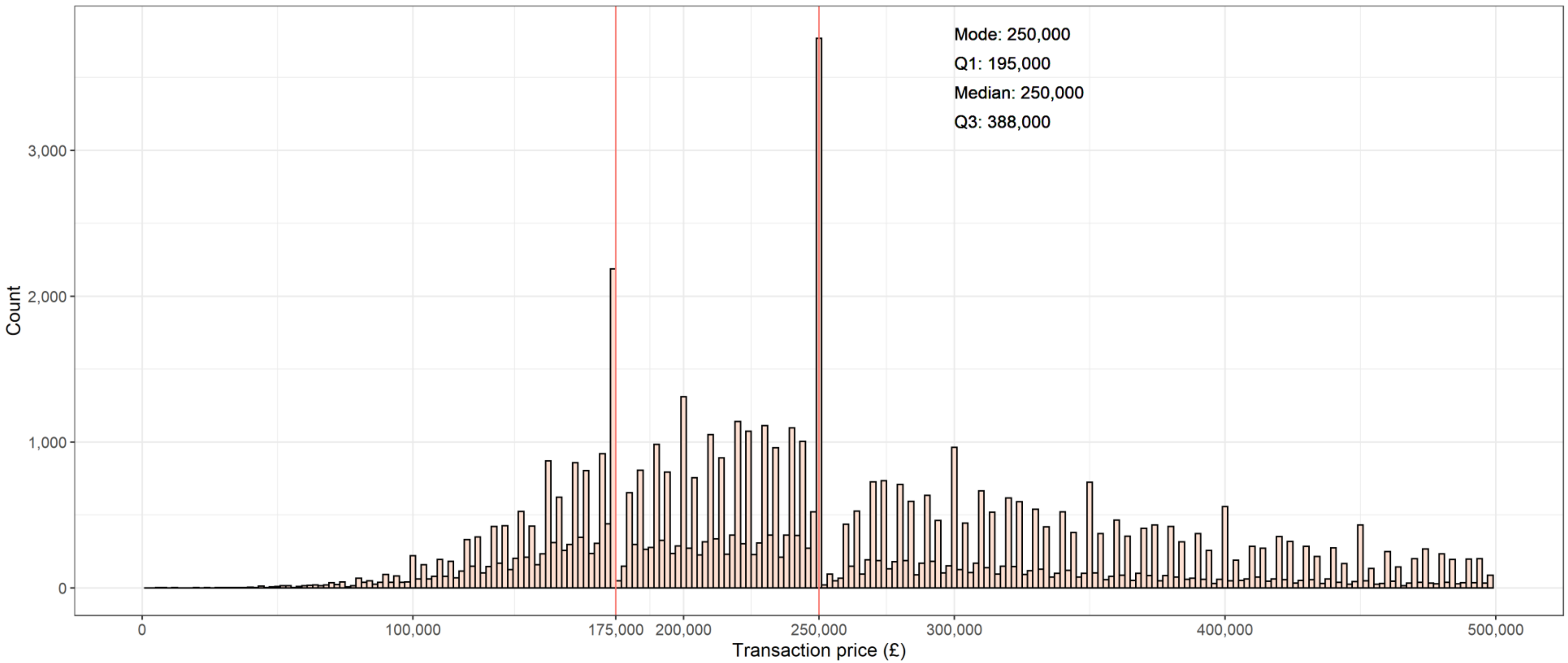


Figure C2.2 The histogram of 2009 TPs in London for the value below £500,000

Appendix C3

Table C3 A list of description of core variables used in this research[38]

Data Sources	Variable	Description	2008	2009	2010	2011	2012	2013	2014	2015	2016
	aacode	Key field									serialanon
	fyear	Fieldwork Year					fYear				-
	tenure8x	Tenure									
	tenure4x	Tenure									
	tenure2x	Tenure									
	GorEHCS	Government office region				gorehcs			-		
generalfs	GorEHS	Region - EHS order				gorehs			gorEHS		
	hhtype6	Household type - 6 categories									
	hhsizex	Number of persons in the household									
	agehrpx	Age of HRP (household reference person) - continuous							SL		
	ager	Report age categories									
	agepartx	Age of partner - continuous							SL		
	emphrpx	Employment status (primary) of HRP									
	emphrp3x	Working status of HRP (primary) - 3 categories									
	empprtx	Employment status (primary) of Partner									
interviewfs	nssech	NS-SEC Socio-economic Classification - HRP							nssech9		

[38] The coloured grid refers to the variable exit in that year. The colored grid with text refers to the new variable names in EHS at this given year.

Data Sources	Variable	Description	2008	2009	2010	2011	2012	2013	2014	2015	2016
	hhempx	Employment status of HRP and partner									
	FreeLeas	Freehold or leasehold				freeleas				freeLeas	
	accomhh	Type of accommodation for household							SL		
	accomhh1	Type of accommodation for household and if not self-contained									
	lenres	Length of residence (years)					lenres2		SL		
	lenresb	Length of residence (catergor)									
	Buypresh	Year HRP bought present accommodation				buypresh			SL		
	ftbuyer	If first-time buyer									
	tenure2	Tenure group 2 (nine categories)									
	tenure3	Tenure group 3 (eight categories)							SL		
	tenure4	Tenure group 4 (five categories)									
	mortwkx	Weekly mortgage payments									
	hhincx	EHS Basic Income (annual net household income (HRP + Partner) including savings)									
	JOINTINCx	Annual gross income of the HRP and partner				jointinc					
	HYEARGRx	Household gross annual income (including income from all adult household members)				hyeargrx					
	ALLincx	Annual gross income of the HRP and partner including income from housing benefit and LHA	-								
	equityh	Equity in home (hybrid version based on mkt value and hhold estimate)			-						
	equityh5	Equity in home (hybrid version based on mkt value and hhold estimate)			-						
	equityr	Equity in home (based on respondent									

Data Sources	Variable	Description	2008	2009	2010	2011	2012	2013	2014	2015	2016
		valuation only)									
	equityr5	Equity in home (based on respondent valuation only)									
	BHCinceq	BHC (Before Housing Costs) equivalised weekly income (modified OECD scale)				bhcinceq					
	AHCinceq	AHC (After Housing Costs) equivalised weekly income (modified OECD scale)				ahcinceq					

Appendix C4

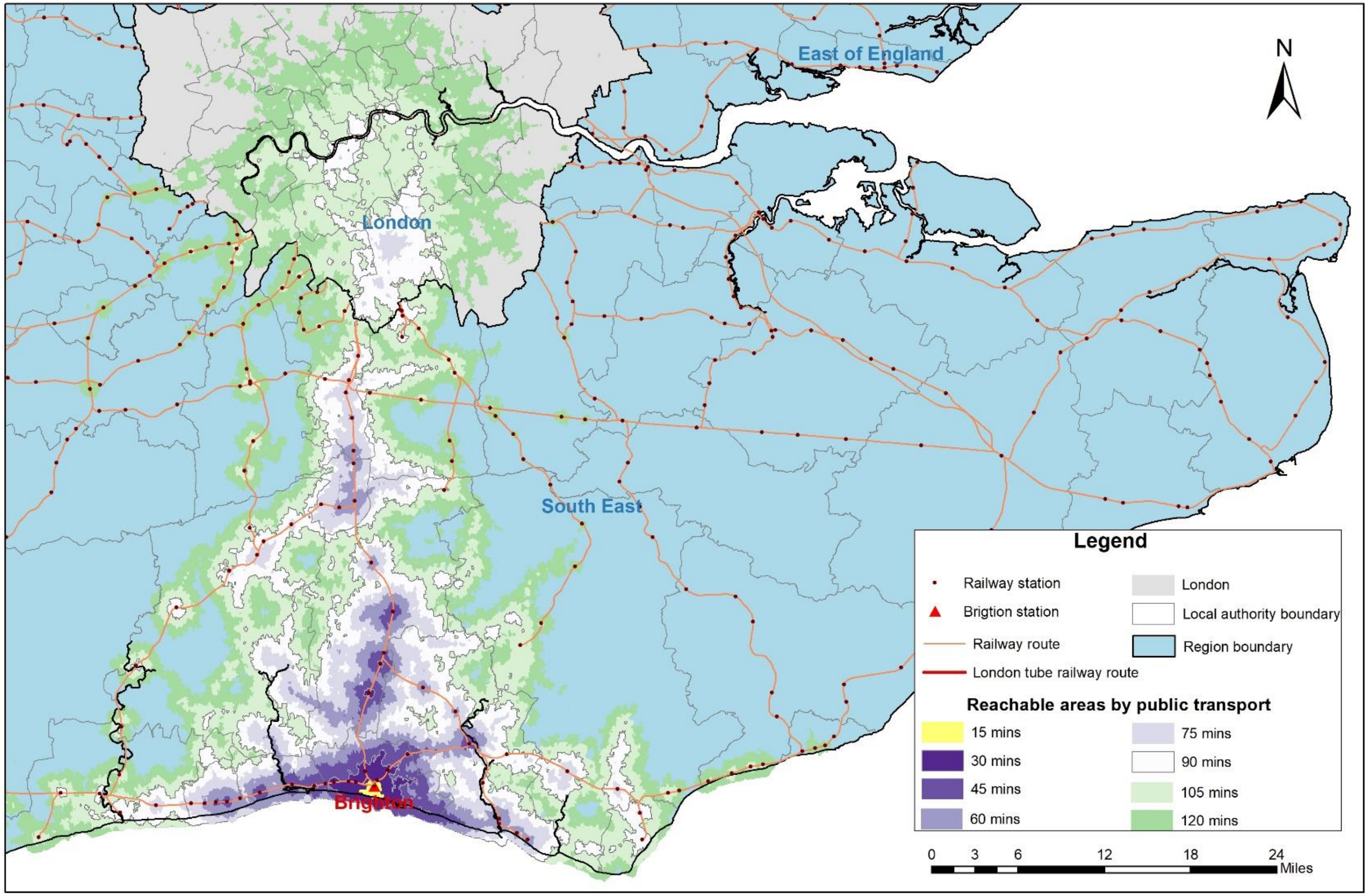


Figure C4 An example of the reachable areas data by public transport at Brighton train station

Appendix C5

Table C5 A summary of the six households characteristic in Scenario B

Scenario ID	**Buyer type**	**Typical household ID**	**Property value (£)**	**Deposit value**		**Monthly mortgage (£)**	**Loan amount**		**Payment for interest rate**	
				(£)	proportion		(£)	proportion	(£)	proportion
Scenario B	First-time mortgage buyers	F1	100,768.10	5,038.41	5.00%	564.5454	95,729.70	95.00%	73,633.92	73.07%
		F2	106,366.30	10,636.63	10.00%	564.5454	95,729.67	90.00%	73,633.92	69.23%
		F3	127,639.60	31,909.90	25.00%	564.5454	95,729.70	75.00%	73,633.92	57.69%
	Non-first-time mortgage buyers in 2009	NF1	124,851	6,242.55	5.00%	699.4682	118,608.45	95.00%	91,231.97	73.07%
		NF2	131,787.20	13,178.72	10.00%	699.4682	118,608.48	90.00%	91,231.97	69.23%
		NF3	158,144.70	39,536.16	25.00%	699.4682	118,608.54	75.00%	91,231.97	57.69%

Appendix C6

Table C6.1 Summaries of affordable size by property type for household C1

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
Wokingham	South East	Flats/maisonette ≈ terraced ≈ semi-detached > detached	57	62	64	64
Christchurch	South West	Flats/maisonettes > semi-detached ≈ terraced > detached	60	72	70	76
Three Rivers	East of England	Flats/maisonettes > semi-detached ≈ terraced > detached	47	56	55	63
North Norfolk	East of England	Flats/maisonettes > semi-detached ≈ terraced > detached	84	96	95	107
Milton Keynes	South East	Flats/maisonettes > terraced > semi-detached > detached	84	93	104	122
Mid Devon	South West	Flats/maisonettes > terraced > semi-detached > detached	77	86	92	105
Hastings	South East	Flats/maisonettes > terraced > semi-detached > detached	84	99	105	114
Rother	South East	Flats/maisonettes > terraced > semi-detached > detached	69	77	83	92
Brentwood	East of England	Flats/maisonettes > terraced > semi-detached > detached	49	59	65	69
North Hertfordshire	East of England	Flats/maisonettes > terraced > semi-detached > detached	59	69	75	82
Maidstone	South East	Flats/maisonettes > terraced > semi-detached > detached	69	76	81	89
Harborough	East Midlands	Flats/maisonettes > terraced > semi-	86	97	105	116

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
		detached > detached				
Oadby and Wigston	East Midlands	Flats/maisonettes > terraced > semi-detached > detached	95	105	115	127
Daventry	East Midlands	Flats/maisonettes > terraced > semi-detached > detached	85	95	106	110
East Northamptonshire	East Midlands	Flats/maisonettes > terraced > semi-detached > detached	98	113	124	130
South Somerset	South West	Flats/maisonettes > terraced > semi-detached > detached	78	91	96	104
West Somerset	South West	Flats/maisonettes > terraced > semi-detached > detached	75	90	96	103
Mid Suffolk	East of England	Flats/maisonettes > terraced > semi-detached > detached	82	93	98	115
Adur	South East	Flats/maisonettes > terraced > semi-detached > detached	63	68	77	81
Worthing	South East	Flats/maisonettes > terraced > semi-detached > detached	67	75	84	88
North Tyneside	North East	Flats/maisonettes > terraced > semi-detached > detached	91	112	125	129
Solihull	West Midlands	Flats/maisonettes > terraced > semi-detached > detached	76	90	96	102
Hackney	London	Flats/maisonettes > terraced ≈ detached ≈ semi-detached	43	42	46	50
Rutland	East Midlands	Flats/maisonettes > terraced ≈ semi-detached > detached	82	96	96	104
Herefordshire,	West Midlands	Flats/maisonettes > terraced ≈ semi-	83	96	97	101

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
County of		detached > detached				
Shropshire	West Midlands	Flats/maisonettes > terraced ≈ semi-detached > detached	83	96	98	102
Aylesbury Vale	South East	Flats/maisonettes > terraced ≈ semi-detached > detached	63	73	74	81
East Cambridgeshire	East of England	Flats/maisonettes > terraced ≈ semi-detached > detached	85	94	94	110
Fenland	East of England	Flats/maisonettes > terraced ≈ semi-detached > detached	117	129	129	136
South Cambridgeshire	East of England	Flats/maisonettes > terraced ≈ semi-detached > detached	65	75	75	87
East Dorset	South West	Flats/maisonettes > terraced ≈ semi-detached > detached	64	69	72	80
North Dorset	South West	Flats/maisonettes > terraced ≈ semi-detached > detached	72	82	83	94
West Dorset	South West	Flats/maisonettes > terraced ≈ semi-detached > detached	67	74	75	83
Wealden	South East	Flats/maisonettes > terraced ≈ semi-detached > detached	62	70	73	81
Braintree	East of England	Flats/maisonettes > terraced ≈ semi-detached > detached	74	84	87	96
Chelmsford	East of England	Flats/maisonettes > terraced ≈ semi-detached > detached	61	70	74	78
Uttlesford	East of England	Flats/maisonettes > terraced ≈ semi-detached > detached	58	67	68	76
Cotswold	South West	Flats/maisonettes > terraced ≈ semi-	57	68	68	76

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
		detached > detached				
Tewkesbury	South West	Flats/maisonettes > terraced ≈ semi-detached > detached	74	84	88	92
Hart	South East	Flats/maisonettes > terraced ≈ semi-detached > detached	56	61	64	69
Winchester	South East	Flats/maisonettes > terraced ≈ semi-detached > detached	53	61	64	73
Broxbourne	East of England	Flats/maisonettes > terraced ≈ semi-detached > detached	60	67	70	74
Dacorum	East of England	Flats/maisonettes > terraced ≈ semi-detached > detached	54	63	67	73
Tonbridge and Malling	South East	Flats/maisonettes > terraced ≈ semi-detached > detached	62	70	73	82
Tunbridge Wells	South East	Flats/maisonettes > terraced ≈ semi-detached > detached	54	64	68	73
South Norfolk	East of England	Flats/maisonettes > terraced ≈ semi-detached > detached	86	94	98	105
South Northamptonshire	East Midlands	Flats/maisonettes > terraced ≈ semi-detached > detached	77	84	88	92
Harrogate	Yorkshire and The Humber	Flats/maisonettes > terraced ≈ semi-detached > detached	68	78	82	87
Ryedale	Yorkshire and The Humber	Flats/maisonettes > terraced ≈ semi-detached > detached	79	89	91	103
South Oxfordshire	South East	Flats/maisonettes > terraced ≈ semi-detached > detached	51	59	60	65
Babergh	East of England	Flats/maisonettes > terraced ≈ semi-	76	87	91	107

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
		detached > detached				
St Edmundsbury	East of England	Flats/maisonettes > terraced ≈ semi-detached > detached	81	92	93	99
Elmbridge	South East	Flats/maisonettes > terraced ≈ semi-detached > detached	41	47	48	55
Guildford	South East	Flats/maisonettes > terraced ≈ semi-detached > detached	47	56	58	63
Mole Valley	South East	Flats/maisonettes > terraced ≈ semi-detached > detached	46	54	54	60
Reigate and Banstead	South East	Flats/maisonettes > terraced ≈ semi-detached > detached	54	62	64	69
Spelthorne	South East	Flats/maisonettes > terraced ≈ semi-detached > detached	54	60	62	67
Tandridge	South East	Flats/maisonettes > terraced ≈ semi-detached > detached	53	63	63	70
Waverley	South East	Flats/maisonettes > terraced ≈ semi-detached > detached	46	54	56	63
Stratford-on-Avon	West Midlands	Flats/maisonettes > terraced ≈ semi-detached > detached	67	75	77	82
Warwick	West Midlands	Flats/maisonettes > terraced ≈ semi-detached > detached	68	79	81	87
Chichester	South East	Flats/maisonettes > terraced ≈ semi-detached > detached	53	62	66	71
Horsham	South East	Flats/maisonettes > terraced ≈ semi-detached > detached	54	64	65	73
Mid Sussex	South East	Flats/maisonettes > terraced ≈ semi-	57	64	66	73

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
		detached > detached				
Malvern Hills	West Midlands	Flats/maisonettes > terraced ≈ semi-detached > detached	78	90	93	106
St Albans	East of England	Flats/maisonettes > terraced ≈ semi-detached > detached	42	50	52	60
East Hertfordshire	East of England	Flats/maisonettes > terraced ≈ semi-detached > detached	53	60	61	70
Barking and Dagenham	London	Flats/maisonettes > terraced ≈ semi-detached > detached	67	81	85	89
Bromley	London	Flats/maisonettes > terraced ≈ semi-detached > detached	52	61	64	69
Ealing	London	Flats/maisonettes > terraced ≈ semi-detached > detached	43	50	52	56
Hounslow	London	Flats/maisonettes > terraced ≈ semi-detached > detached	49	54	56	61
Southwark	London	Flats/maisonettes > terraced ≈ semi-detached > detached	39	44	47	51
Wandsworth	London	Flats/maisonettes > terraced ≈ semi-detached > detached	29	34	37	42
Brighton and Hove	South East	Flats/maisonettes > terraced ≈ semi-detached ≈ detached	53	57	60	66
Broadland	East of England	Flats/maisonettes > terraced ≈ semi-detached ≈ detached	89	93	95	100
Camden	London	Flats/maisonettes > terraced ≈ semi-detached ≈ detached	19	23	26	30
Hammersmith and	London	Flats/maisonettes > terraced ≈ semi-	28	29	29	35

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
Fulham		detached ≈ detached				
Kensington and Chelsea	London	Flats/maisonettes > terraced ≈ semi-detached ≈ detached	14	15	15	22
Kingston upon Thames	London	Flats/maisonettes > terraced ≈ semi-detached ≈ detached	44	48	50	54
Richmond upon Thames	London	Flats/maisonettes > terraced ≈ semi-detached ≈ detached	30	34	36	43
Westminster	London	Flats/maisonettes > terraced ≈ semi-detached ≈ detached	16	19	19	25
Cambridge	East of England	Flats/maisonettes ≈ semi-detached ≈ terraced > detached	51	58	57	60
Suffolk Coastal	East of England	Flats/maisonettes ≈ semi-detached ≈ terraced > detached	78	88	87	91
Surrey Heath	South East	Flats/maisonettes ≈ semi-detached ≈ terraced > detached	58	65	64	69
Bromsgrove	West Midlands	Flats/maisonettes ≈ semi-detached ≈ terraced > detached	76	89	88	91
Thurrock	East of England	Flats/maisonettes ≈ terraced > semi-detached > detached	72	80	91	92
Teignbridge	South West	Flats/maisonettes ≈ terraced > semi-detached > detached	72	82	87	87
Eastbourne	South East	Flats/maisonettes ≈ terraced > semi-detached > detached	77	84	91	94
Harlow	East of England	Flats/maisonettes ≈ terraced > semi-detached > detached	74	81	93	94
Cheltenham	South West	Flats/maisonettes ≈ terraced > semi-	71	79	85	87

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
		detached > detached				
Basingstoke and Deane	South East	Flats/maisonettes ≈ terraced > semi-detached > detached	65	73	81	82
Ashford	South East	Flats/maisonettes ≈ terraced > semi-detached > detached	73	83	89	92
Dartford	South East	Flats/maisonettes ≈ terraced > semi-detached > detached	66	79	84	86
Dover	South East	Flats/maisonettes ≈ terraced > semi-detached > detached	83	100	108	110
Gravesham	South East	Flats/maisonettes ≈ terraced > semi-detached > detached	69	84	93	93
Charnwood	East Midlands	Flats/maisonettes ≈ terraced > semi-detached > detached	90	103	109	111
Breckland	East of England	Flats/maisonettes ≈ terraced > semi-detached > detached	98	107	116	116
Richmondshire	Yorkshire and The Humber	Flats/maisonettes ≈ terraced > semi-detached > detached	83	94	99	99
Arun	South East	Flats/maisonettes ≈ terraced > semi-detached > detached	69	76	83	85
Crawley	South East	Flats/maisonettes ≈ terraced > semi-detached > detached	70	76	83	83
Welwyn Hatfield	East of England	Flats/maisonettes ≈ terraced > semi-detached > detached	54	64	69	71
Newcastle upon Tyne	North East	Flats/maisonettes ≈ terraced > semi-detached > detached	89	118	127	127
Birmingham	West Midlands	Flats/maisonettes ≈ terraced > semi-	94	116	128	131

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
		detached > detached				
Coventry	West Midlands	Flats/maisonettes ≈ terraced > semi-detached > detached	96	120	133	134
Bradford	Yorkshire and The Humber	Flats/maisonettes ≈ terraced > semi-detached > detached	100	119	139	139
Gateshead	North East	Flats/maisonettes ≈ terraced > semi-detached > detached	95	120	138	140
Haringey	London	Flats/maisonettes ≈ terraced > semi-detached > detached	38	47	53	56
Havering	London	Flats/maisonettes ≈ terraced > semi-detached > detached	59	70	75	76
Slough	South East	Flats/maisonettes ≈ terraced > semi-detached ≈ detached	66	70	75	76
Fareham	South East	Flats/maisonettes ≈ terraced > semi-detached ≈ detached	71	75	82	85
Bracknell Forest	South East	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	62	67	71	72
West Berkshire	South East	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	61	68	70	72
Windsor and Maidenhead	South East	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	48	56	56	57
Wiltshire	South West	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	75	84	87	88
Bedford	East of England	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	85	95	99	102
Central	East of England	Flats/maisonettes ≈ terraced ≈ semi-	77	85	88	89

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
Bedfordshire		detached > detached				
Chiltern	South East	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	46	54	58	60
South Bucks	South East	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	48	55	58	60
Wycombe	South East	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	53	64	65	66
East Devon	South West	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	65	76	80	82
Exeter	South West	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	71	78	80	80
West Devon	South West	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	76	92	94	95
Lewes	South East	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	63	69	73	74
Colchester	East of England	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	79	88	92	94
Epping Forest	East of England	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	50	59	63	63
Stroud	South West	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	70	83	86	89
East Hampshire	South East	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	60	68	72	75
Hertsmere	East of England	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	47	56	59	61
Sevenoaks	South East	Flats/maisonettes ≈ terraced ≈ semi-	52	64	68	69

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
		detached > detached				
King's Lynn and West Norfolk	East of England	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	100	112	113	115
Norwich	East of England	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	88	102	105	107
Rushcliffe	East Midlands	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	84	97	100	102
Cherwell	South East	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	67	75	77	80
Oxford	South East	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	47	54	56	56
Vale of White Horse	South East	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	59	67	68	71
West Oxfordshire	South East	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	60	65	66	68
Taunton Deane	South West	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	79	91	94	96
Lichfield	West Midlands	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	81	96	100	100
Epsom and Ewell	South East	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	49	55	59	61
Runnymede	South East	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	50	57	58	60
Woking	South East	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	50	58	59	61
Worcester	West Midlands	Flats/maisonettes ≈ terraced ≈ semi-	84	96	99	100

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
		detached > detached				
Wychavon	West Midlands	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	76	86	90	90
Barnet	London	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	38	46	49	52
Bexley	London	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	64	74	78	78
Brent	London	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	46	51	54	56
Croydon	London	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	61	71	75	76
Enfield	London	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	53	61	63	66
Greenwich	London	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	52	68	72	72
Harrow	London	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	48	54	56	57
Hillingdon	London	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	53	59	61	62
Lambeth	London	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	39	44	48	49
Lewisham	London	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	54	63	65	67
Merton	London	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	38	49	51	54
Redbridge	London	Flats/maisonettes ≈ terraced ≈ semi-	54	62	65	67

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
		detached > detached				
Sutton	London	Flats/maisonettes ≈ terraced ≈ semi-detached > detached	55	62	64	67
Canterbury	South East	Flats/maisonettes ≈ terraced ≈ semi-detached ≈ detached	74	78	82	85
Islington	London	Flats/maisonettes ≈ terraced ≈ semi-detached ≈ detached	31	32	34	37
Newham	London	Flats/maisonettes ≈ terraced ≈ semi-detached ≈ detached	70	74	77	77
Tower Hamlets	London	Semi-detached ≈ terraced≈ flats/maisonettes> detached	36	46	45	46
Bournemouth	South West	Semi-detached ≈ flats/maisonettes ≈ terraced > detached	74	84	82	83
Poole	South West	Semi-detached ≈ flats/maisonettes ≈ terraced > detached	67	78	75	77
South Hams	South West	Semi-detached ≈ terraced ≈ flats/maisonettes > detached	60	73	71	71
Maldon	East of England	Semi-detached ≈ terraced ≈ flats/maisonettes > detached	73	82	81	81
Eastleigh	South East	Semi-detached ≈ terraced ≈ flats/maisonettes > detached	68	79	78	78
Hartlepool	North East	Terraced > flats/maisonettes > semi-detached > detached	114	144	189	151
Redcar and Cleveland	North East	Terraced > flats/maisonettes > semi-detached > detached	106	127	166	134
Stockton-on-Tees	North East	Terraced > flats/maisonettes > semi-	107	129	157	137

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
		detached > detached				
Halton	North West	Terraced > flats/maisonettes > semi-detached > detached	107	126	157	146
Blackburn with Darwen	North West	Terraced > flats/maisonettes > semi-detached > detached	109	125	181	145
East Riding of Yorkshire	Yorkshire and The Humber	Terraced > flats/maisonettes > semi-detached > detached	99	112	125	115
North East Lincolnshire	Yorkshire and The Humber	Terraced > flats/maisonettes > semi-detached > detached	123	144	179	166
North Lincolnshire	Yorkshire and The Humber	Terraced > flats/maisonettes > semi-detached > detached	122	148	173	159
York	Yorkshire and The Humber	Terraced > flats/maisonettes > semi-detached > detached	73	82	87	83
Derby	East Midlands	Terraced > flats/maisonettes > semi-detached > detached	107	127	136	131
Leicester	East Midlands	Terraced > flats/maisonettes > semi-detached > detached	105	128	137	131
Telford and Wrekin	West Midlands	Terraced > flats/maisonettes > semi-detached > detached	102	116	129	124
Torbay	South West	Terraced > flats/maisonettes > semi-detached > detached	81	94	106	97
Peterborough	East of England	Terraced > flats/maisonettes > semi-detached > detached	106	121	134	125
Southend-on-Sea	East of England	Terraced > flats/maisonettes > semi-detached > detached	75	86	96	90
Cheshire East	North West	Terraced > flats/maisonettes > semi-	83	97	110	100

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
		detached > detached				
Cheshire West and Chester	North West	Terraced > flats/maisonettes > semi-detached > detached	87	104	116	107
Northumberland	North East	Terraced > flats/maisonettes > semi-detached > detached	93	117	131	124
Barrow-in-Furness	North West	Terraced > flats/maisonettes > semi-detached > detached	98	120	158	145
Carlisle	North West	Terraced > flats/maisonettes > semi-detached > detached	101	125	139	129
Eden	North West	Terraced > flats/maisonettes > semi-detached > detached	83	97	102	98
Amber Valley	East Midlands	Terraced > flats/maisonettes > semi-detached > detached	100	129	142	138
Bolsover	East Midlands	Terraced > flats/maisonettes > semi-detached > detached	116	159	207	163
Erewash	East Midlands	Terraced > flats/maisonettes > semi-detached > detached	106	135	150	140
North East Derbyshire	East Midlands	Terraced > flats/maisonettes > semi-detached > detached	92	117	130	123
South Derbyshire	East Midlands	Terraced > flats/maisonettes > semi-detached > detached	97	115	126	119
Torridge	South West	Terraced > flats/maisonettes > semi-detached > detached	78	90	99	95
Basildon	East of England	Terraced > flats/maisonettes > semi-detached > detached	70	79	92	86
Swale	South East	Terraced > flats/maisonettes > semi-	87	98	111	107

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
		detached > detached				
Burnley	North West	Terraced > flats/maisonettes > semi-detached > detached	111	126	202	155
Hyndburn	North West	Terraced > flats/maisonettes > semi-detached > detached	103	119	205	163
Lancaster	North West	Terraced > flats/maisonettes > semi-detached > detached	89	106	121	114
Pendle	North West	Terraced > flats/maisonettes > semi-detached > detached	104	131	194	133
Preston	North West	Terraced > flats/maisonettes > semi-detached > detached	103	115	137	130
Ribble Valley	North West	Terraced > flats/maisonettes > semi-detached > detached	74	87	105	91
Rossendale	North West	Terraced > flats/maisonettes > semi-detached > detached	101	125	161	129
South Ribble	North West	Terraced > flats/maisonettes > semi-detached > detached	92	103	126	112
West Lancashire	North West	Terraced > flats/maisonettes > semi-detached > detached	96	110	130	124
Wyre	North West	Terraced > flats/maisonettes > semi-detached > detached	93	106	124	112
Hinckley and Bosworth	East Midlands	Terraced > flats/maisonettes > semi-detached > detached	93	107	118	112
North West Leicestershire	East Midlands	Terraced > flats/maisonettes > semi-detached > detached	98	118	130	124
East Lindsey	East Midlands	Terraced > flats/maisonettes > semi-	111	125	150	140

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
		detached > detached				
North Kesteven	East Midlands	Terraced > flats/maisonettes > semi-detached > detached	109	118	124	119
South Kesteven	East Midlands	Terraced > flats/maisonettes > semi-detached > detached	102	117	129	122
Great Yarmouth	East of England	Terraced > flats/maisonettes > semi-detached > detached	108	116	129	122
Kettering	East Midlands	Terraced > flats/maisonettes > semi-detached > detached	105	122	137	124
Wellingborough	East Midlands	Terraced > flats/maisonettes > semi-detached > detached	103	120	144	132
Craven	Yorkshire and The Humber	Terraced > flats/maisonettes > semi-detached > detached	75	86	96	87
Scarborough	Yorkshire and The Humber	Terraced > flats/maisonettes > semi-detached > detached	86	103	108	104
Selby	Yorkshire and The Humber	Terraced > flats/maisonettes > semi-detached > detached	94	105	117	110
Broxtowe	East Midlands	Terraced > flats/maisonettes > semi-detached > detached	101	125	139	128
Gedling	East Midlands	Terraced > flats/maisonettes > semi-detached > detached	103	125	133	126
Sedgemoor	South West	Terraced > flats/maisonettes > semi-detached > detached	87	93	109	105
Newcastle-under-Lyme	West Midlands	Terraced > flats/maisonettes > semi-detached > detached	106	132	155	138
Tamworth	West Midlands	Terraced > flats/maisonettes > semi-	95	111	134	116

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
		detached > detached				
Forest Heath	East of England	Terraced > flats/maisonettes > semi-detached > detached	89	102	119	113
Waveney	East of England	Terraced > flats/maisonettes > semi-detached > detached	95	106	118	112
Nuneaton and Bedworth	West Midlands	Terraced > flats/maisonettes > semi-detached > detached	98	119	136	127
Rugby	West Midlands	Terraced > flats/maisonettes > semi-detached > detached	91	107	121	114
Redditch	West Midlands	Terraced > flats/maisonettes > semi-detached > detached	91	103	120	114
Bury	North West	Terraced > flats/maisonettes > semi-detached > detached	94	116	141	121
Oldham	North West	Terraced > flats/maisonettes > semi-detached > detached	96	117	153	130
Rochdale	North West	Terraced > flats/maisonettes > semi-detached > detached	99	123	157	144
Stockport	North West	Terraced > flats/maisonettes > semi-detached > detached	84	98	112	105
Tameside	North West	Terraced > flats/maisonettes > semi-detached > detached	98	122	150	129
Trafford	North West	Terraced > flats/maisonettes > semi-detached > detached	77	93	105	99
Wigan	North West	Terraced > flats/maisonettes > semi-detached > detached	103	127	174	135
Knowsley	North West	Terraced > flats/maisonettes > semi-	113	139	171	152

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
		detached > detached				
Liverpool	North West	Terraced > flats/maisonettes > semi-detached > detached	113	138	164	144
St. Helens	North West	Terraced > flats/maisonettes > semi-detached > detached	100	125	164	130
Sunderland	North East	Terraced > flats/maisonettes > semi-detached > detached	105	134	162	144
Sandwell	West Midlands	Terraced > flats/maisonettes > semi-detached > detached	103	132	147	142
Calderdale	Yorkshire and The Humber	Terraced > flats/maisonettes > semi-detached > detached	95	114	139	127
Kirklees	Yorkshire and The Humber	Terraced > flats/maisonettes > semi-detached > detached	97	117	137	133
Castle Point	East of England	Terraced > flats/maisonettes > semi-detached ≈ detached	76	77	88	82
Salford	North West	Terraced > flats/maisonettes ≈ semi-detached > detached	106	130	171	131
South Staffordshire	West Midlands	Terraced > flats/maisonettes ≈ semi-detached > detached	82	96	103	99
North Warwickshire	West Midlands	Terraced > flats/maisonettes ≈ semi-detached > detached	85	106	115	109
Stevenage	East of England	Terraced > flats/maisonettes ≈ semi-detached > detached	73	85	96	89
Bolton	North West	Terraced > flats/maisonettes ≈ semi-detached > detached	102	128	164	132
Sheffield	Yorkshire and The	Terraced > flats/maisonettes ≈ semi-	97	119	136	122

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
	Humber	detached > detached				
Dudley	West Midlands	Terraced > flats/maisonettes ≈ semi-detached > detached	95	115	125	116
Walsall	West Midlands	Terraced > flats/maisonettes ≈ semi-detached > detached	104	128	141	132
Wolverhampton	West Midlands	Terraced > flats/maisonettes ≈ semi-detached > detached	105	138	154	139
Leeds	Yorkshire and The Humber	Terraced > flats/maisonettes ≈ semi-detached > detached	90	107	116	111
Middlesbrough	North East	Terraced > semi-detached > flats/maisonettes > detached	112	140	158	130
Darlington	North East	Terraced > semi-detached > flats/maisonettes > detached	109	125	154	114
Warrington	North West	Terraced > semi-detached > flats/maisonettes > detached	91	111	127	110
Blackpool	North West	Terraced > semi-detached > flats/maisonettes > detached	119	150	171	132
Nottingham	East Midlands	Terraced > semi-detached > flats/maisonettes > detached	117	151	168	140
Stoke-on-Trent	West Midlands	Terraced > semi-detached > flats/maisonettes > detached	115	152	191	150
Bristol, City of	South West	Terraced > semi-detached > flats/maisonettes > detached	80	96	101	90
Plymouth	South West	Terraced > semi-detached > flats/maisonettes > detached	89	108	116	106
Reading	South East	Terraced > semi-detached >	66	73	79	72

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
		flats/maisonettes > detached				
Southampton	South East	Terraced > semi-detached > flats/maisonettes > detached	82	92	97	90
County Durham	North East	Terraced > semi-detached > flats/maisonettes > detached	110	145	181	138
Allerdale	North West	Terraced > semi-detached > flats/maisonettes > detached	91	118	139	117
Copeland	North West	Terraced > semi-detached > flats/maisonettes > detached	110	144	175	137
Chesterfield	East Midlands	Terraced > semi-detached > flats/maisonettes > detached	99	133	149	128
High Peak	East Midlands	Terraced > semi-detached > flats/maisonettes > detached	87	105	116	103
North Devon	South West	Terraced > semi-detached > flats/maisonettes > detached	75	85	97	83
Chorley	North West	Terraced > semi-detached > flats/maisonettes > detached	90	103	131	101
Fylde	North West	Terraced > semi-detached > flats/maisonettes > detached	85	103	118	98
Boston	East Midlands	Terraced > semi-detached > flats/maisonettes > detached	125	143	149	141
Lincoln	East Midlands	Terraced > semi-detached > flats/maisonettes > detached	107	125	144	122
Ashfield	East Midlands	Terraced > semi-detached > flats/maisonettes > detached	120	160	192	154
Bassetlaw	East Midlands	Terraced > semi-detached >	115	150	170	133

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
		flats/maisonettes > detached				
Mansfield	East Midlands	Terraced > semi-detached > flats/maisonettes > detached	121	163	198	153
Barnsley	Yorkshire and The Humber	Terraced > semi-detached > flats/maisonettes > detached	110	145	168	138
Doncaster	Yorkshire and The Humber	Terraced > semi-detached > flats/maisonettes > detached	113	147	165	131
Rotherham	Yorkshire and The Humber	Terraced > semi-detached > flats/maisonettes > detached	111	146	167	137
South Holland	East Midlands	Terraced > semi-detached > flats/maisonettes ≈ detached	123	136	141	124
East Staffordshire	West Midlands	Terraced > semi-detached ≈ flats/maisonettes > detached	97	116	130	116
Staffordshire Moorlands	West Midlands	Terraced > semi-detached ≈ flats/maisonettes > detached	94	117	125	115
Manchester	North West	Terraced > semi-detached ≈ flats/maisonettes > detached	104	124	141	122
Sefton	North West	Terraced > semi-detached ≈ flats/maisonettes > detached	103	123	136	120
Wirral	North West	Terraced > semi-detached ≈ flats/maisonettes > detached	100	126	139	125
Wakefield	Yorkshire and The Humber	Terraced > semi-detached ≈ flats/maisonettes > detached	97	125	147	123
Kingston upon Hull, City of	Yorkshire and The Humber	Terraced > semi-detached=flats/maisonettes > detached	122	153	190	153
South Lakeland	North West	Terraced > semi-	65	76	81	76

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
		detached=flats/maisonettes > detached				
City of London	London	terraced ≈ flats/maisonettes	NA	NA	32	36
Bath and North East Somerset	South West	Terraced ≈ flats/maisonettes > semi-detached > detached	65	75	80	78
North Somerset	South West	Terraced ≈ flats/maisonettes > semi-detached > detached	81	89	96	93
South Gloucestershire	South West	Terraced ≈ flats/maisonettes > semi-detached > detached	77	86	92	91
Swindon	South West	Terraced ≈ flats/maisonettes > semi-detached > detached	93	104	114	111
Luton	East of England	Terraced ≈ flats/maisonettes > semi-detached > detached	89	98	105	103
Medway	South East	Terraced ≈ flats/maisonettes > semi-detached > detached	83	96	109	107
Portsmouth	South East	Terraced ≈ flats/maisonettes > semi-detached > detached	81	96	104	102
Tendring	East of England	Terraced ≈ flats/maisonettes > semi-detached > detached	91	101	111	110
Forest of Dean	South West	Terraced ≈ flats/maisonettes > semi-detached > detached	87	104	109	108
Gloucester	South West	Terraced ≈ flats/maisonettes > semi-detached > detached	96	107	114	111
Gosport	South East	Terraced ≈ flats/maisonettes > semi-detached > detached	83	92	107	105
Havant	South East	Terraced ≈ flats/maisonettes > semi-detached > detached	74	87	96	94

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
Test Valley	South East	Terraced ≈ flats/maisonettes > semi-detached > detached	63	74	79	78
Shepway	South East	Terraced ≈ flats/maisonettes > semi-detached > detached	81	89	100	98
Thanet	South East	Terraced ≈ flats/maisonettes > semi-detached > detached	89	101	115	112
Blaby	East Midlands	Terraced ≈ flats/maisonettes > semi-detached > detached	95	106	115	112
Melton	East Midlands	Terraced ≈ flats/maisonettes > semi-detached > detached	89	106	117	116
West Lindsey	East Midlands	Terraced ≈ flats/maisonettes > semi-detached > detached	114	133	142	139
Corby	East Midlands	Terraced ≈ flats/maisonettes > semi-detached > detached	115	142	161	159
Northampton	East Midlands	Terraced ≈ flats/maisonettes > semi-detached > detached	97	105	127	126
Newark and Sherwood	East Midlands	Terraced ≈ flats/maisonettes > semi-detached > detached	100	124	133	132
Mendip	South West	Terraced ≈ flats/maisonettes > semi-detached > detached	77	85	91	89
Ipswich	East of England	Terraced ≈ flats/maisonettes > semi-detached > detached	96	113	124	123
Wyre Forest	West Midlands	Terraced ≈ flats/maisonettes > semi-detached > detached	90	105	112	110
South Tyneside	North East	Terraced ≈ flats/maisonettes > semi-detached > detached	94	124	151	148

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
Waltham Forest	London	Terraced ≈ flats/maisonettes > semi-detached > detached	58	63	70	69
New Forest	South East	Terraced ≈ flats/maisonettes ≈ semi-detached > detached	60	70	73	72
Rushmoor	South East	Terraced ≈ flats/maisonettes ≈ semi-detached > detached	68	76	80	79
Watford	East of England	Terraced ≈ flats/maisonettes ≈ semi-detached > detached	58	64	67	66
Rochford	East of England	Terraced ≈ flats/maisonettes ≈ semi-detached ≈ detached	71	75	79	78
Cornwall	South West	Terraced ≈ semi-detached > flats/maisonettes > detached	74	85	88	82
Derbyshire Dales	East Midlands	Terraced ≈ semi-detached > flats/maisonettes > detached	73	89	91	87
Weymouth and Portland	South West	Terraced ≈ semi-detached > flats/maisonettes > detached	75	89	92	85
Hambleton	Yorkshire and The Humber	Terraced ≈ semi-detached > flats/maisonettes > detached	78	87	91	83
Isles of Scilly	South West	Terraced ≈ semi-detached > flats/maisonettes ≈ detached	62	67	71	66
Isle of Wight	South East	Terraced ≈ semi-detached ≈ flats/maisonettes > detached	74	92	93	92
Huntingdonshire	East of England	Terraced ≈ semi-detached ≈ flats/maisonettes > detached	88	98	99	98
Purbeck	South West	Terraced ≈ semi-detached ≈ flats/maisonettes > detached	62	74	75	73

LAs	Region	Rank of affordable size	Affordable property size (m^2)			
			Detached	Semi-detached	Terraced	Flats/maisonettes
Cannock Chase	West Midlands	Terraced ≈ semi-detached ≈ flats/maisonettes > detached	104	124	128	120
Stafford	West Midlands	Terraced ≈ semi-detached ≈ flats/maisonettes > detached	93	107	110	106

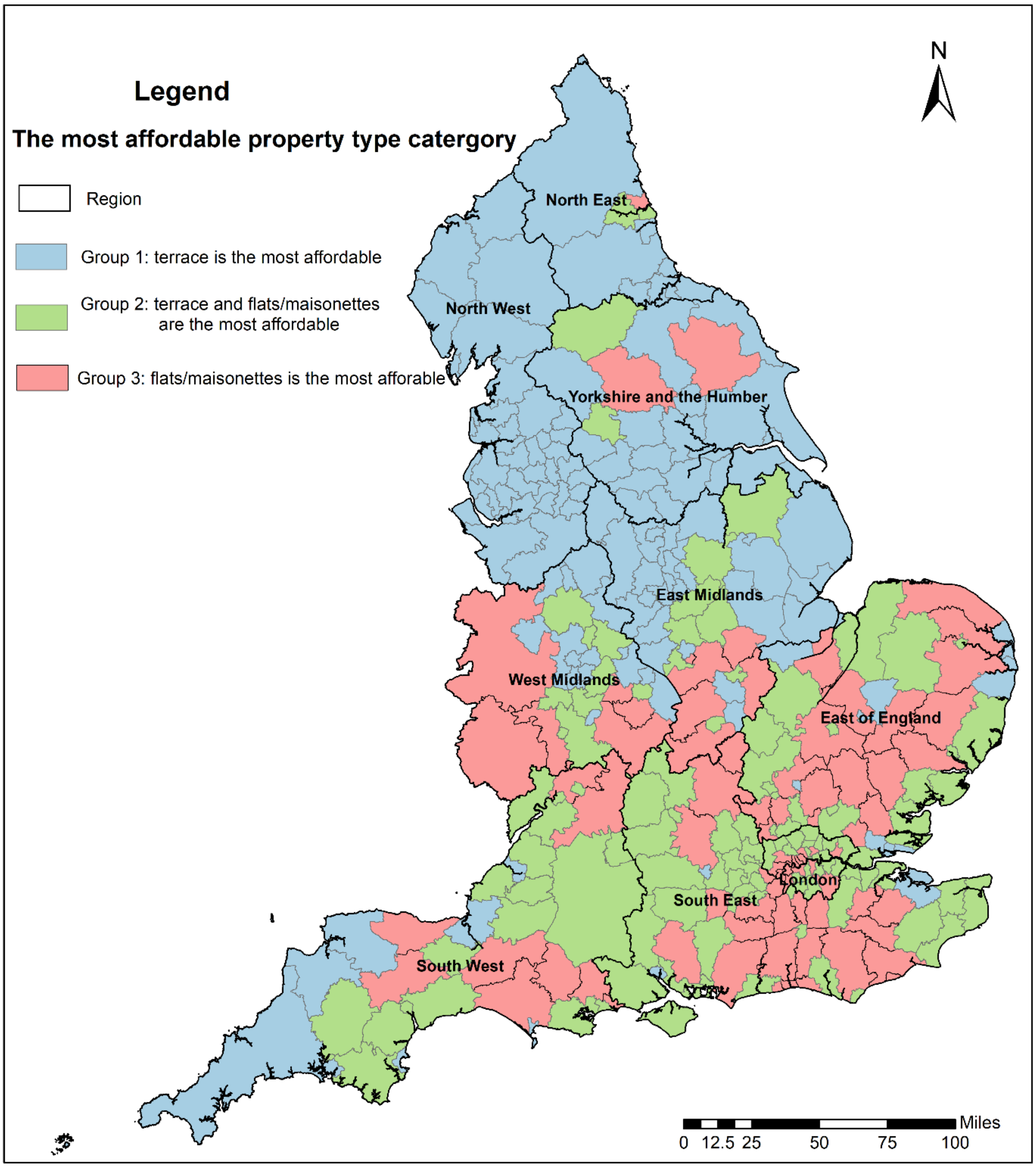


Figure C6 Spatial map of the affordable size order in property types in England at LA level

Table C6.2 The order of affordable property size among the four property types at LAs scale

Group No.	Secondary category	Region	No.	LA
Group 1	Terraced > flats/maisonettes > semi-detached > detached	East Midlands	1	Derby
			2	Leicester
			3	Amber Valley
			4	Bolsover
			5	Erewash
			6	North East Derbyshire
			7	South Derbyshire
			8	Hinckley and Bosworth
			9	North West Leicestershire
			10	East Lindsey
			11	North Kesteven
			12	South Kesteven
			13	Kettering
			14	Wellingborough
			15	Broxtowe
			16	Gedling
		East of England	17	Peterborough
			18	Southend-on-Sea
			19	Basildon
			20	Great Yarmouth
			21	Forest Heath
			22	Waveney
		North East	23	Hartlepool
			24	Redcar and Cleveland
			25	Stockton-on-Tees
			26	Northumberland
			27	Sunderland
		North West	28	Halton
			29	Blackburn with Darwen
			30	Cheshire East
			31	Cheshire West and Chester
			32	Barrow-in-Furness
			33	Carlisle
			34	Eden
			35	Burnley
			36	Hyndburn
			37	Lancaster

Group No.	Secondary category	Region	No.	LA
			38	Pendle
			39	Preston
			40	Ribble Valley
			41	Rossendale
			42	South Ribble
			43	West Lancashire
			44	Wyre
			45	Bury
			46	Oldham
			47	Rochdale
			48	Stockport
			49	Tameside
			50	Trafford
			51	Wigan
			52	Knowsley
			53	Liverpool
			54	St. Helens
		South East	55	Swale
		South West	56	Torbay
			57	Torridge
			58	Sedgemoor
		West Midlands	59	Telford and Wrekin
			60	Newcastle-under-Lyme
			61	Tamworth
			62	Nuneaton and Bedworth
			63	Rugby
			64	Redditch
			65	Sandwell
		Yorkshire and The Humber	66	East Riding of Yorkshire
			67	North East Lincolnshire
			68	North Lincolnshire
			69	York
			70	Craven
			71	Scarborough
			72	Selby
			73	Calderdale
			74	Kirklees
	Terraced > semi-detached > flats/maisonettes >	East Midlands	1	Nottingham
			2	Chesterfield
			3	High Peak

Group No.	Secondary category	Region	No.	LA
	detached		4	Boston
			5	Lincoln
			6	Ashfield
			7	Bassetlaw
			8	Mansfield
		North East	9	Middlesbrough
			10	Darlington
			11	County Durham
		North West	12	Warrington
			13	Blackpool
			14	Allerdale
			15	Copeland
			16	Chorley
			17	Fylde
		South East	18	Reading
			19	Southampton
		South West	20	Bristol, City of
			21	Plymouth
			22	North Devon
		West Midlands	23	Stoke-on-Trent
		Yorkshire and The Humber	24	Barnsley
			25	Doncaster
			26	Rotherham
	Terraced > flats/maisonettes ≈ semi-detached > detached	East of England	1	Stevenage
		North West	2	Salford
			3	Bolton
			4	Manchester
			5	Sefton
			6	Wirral
			7	South Lakeland
		West Midlands	8	South Staffordshire
			9	North Warwickshire
			10	Dudley
			11	Walsall
			12	Wolverhampton
			13	East Staffordshire
			14	Staffordshire Moorlands
		Yorkshire and	15	Sheffield
			16	Leeds

Group No.	Secondary category	Region	No.	LA
		The Humber	17	Wakefield
			18	Kingston upon Hull, City of
	Terraced ≈ semi-detached > flats/maisonettes > detached	East Midlands	1	Derbyshire Dales
		South West	2	Cornwall
			3	Weymouth and Portland
		Yorkshire and The Humber	4	Hambleton
	Terraced > flats/maisonettes > semi-detached ≈ detached	East of England	1	Castle Point
	Terraced > semi-detached > flats/maisonettes ≈ detached	East Midlands	1	South Holland
	Terraced ≈ semi-detached > flats/maisonettes ≈ detached	South West	1	Isles of Scilly
Group 2	Flats/maisonette ≈ terraced ≈ semi-detached > detached	East Midlands	1	Rushcliffe
		East of England	2	Cambridge
			3	Suffolk Coastal
			4	Bedford
			5	Central Bedfordshire
			6	Colchester
			7	Epping Forest
			8	Hertsmere
			9	King's Lynn and West Norfolk
			10	Norwich
			11	Maldon
			12	Watford
			13	Huntingdonshire
		London	14	Barnet
			15	Bexley
			16	Brent
			17	Croydon
			18	Enfield
			19	Greenwich
			20	Harrow
			21	Hillingdon

Group No.	Secondary category	Region	No.	LA
			22	Lambeth
			23	Lewisham
			24	Merton
			25	Redbridge
			26	Sutton
			27	Tower Hamlets
		South East	28	Wokingham
			29	Surrey Heath
			30	Bracknell Forest
			31	West Berkshire
			32	Windsor and Maidenhead
			33	Chiltern
			34	South Bucks
			35	Wycombe
			36	Lewes
			37	East Hampshire
			38	Sevenoaks
			39	Cherwell
			40	Oxford
			41	Vale of White Horse
			42	West Oxfordshire
			43	Epsom and Ewell
			44	Runnymede
			45	Woking
			46	Eastleigh
			47	New Forest
			48	Rushmoor
			49	Isle of Wight
		South West	50	Wiltshire
			51	East Devon
			52	Exeter
			53	West Devon
			54	Stroud
			55	Taunton Deane
			56	Bournemouth
			57	Poole
			58	South Hams
			59	Purbeck
		West Midlands	60	Bromsgrove
			61	Lichfield
			62	Worcester

Group No.	Secondary category	Region	No.	LA
			63	Wychavon
			64	Cannock Chase
			65	Stafford
		East Midlands	1	Charnwood
			2	Blaby
			3	Melton
			4	West Lindsey
			5	Corby
			6	Northampton
			7	Newark and Sherwood
		East of England	8	Thurrock
			9	Harlow
			10	Breckland
			11	Welwyn Hatfield
			12	Luton
			13	Tendring
			14	Ipswich
		London	15	Haringey
			16	Havering
			17	Waltham Forest
		North East	18	Newcastle upon Tyne
			19	Gateshead
			20	South Tyneside
		South East	21	Eastbourne
			22	Basingstoke and Deane
			23	Ashford
			24	Dartford
			25	Dover
			26	Gravesham
			27	Arun
			28	Crawley
			29	Medway
			30	Portsmouth
			31	Gosport
			32	Havant
			33	Test Valley
			34	Shepway
			35	Thanet
	Flats/maisonette ≈ terraced > semi-detached > detached	South West	36	Teignbridge
			37	Cheltenham
			38	Bath and North East

Group No.	Secondary category	Region	No.	LA
				Somerset
			39	North Somerset
			40	South Gloucestershire
			41	Swindon
			42	Forest of Dean
			43	Gloucester
			44	Mendip
		West Midlands	45	Birmingham
			46	Coventry
			47	Wyre Forest
		Yorkshire and The Humber	48	Richmondshire
			49	Bradford
	Flats/maisonettes ≈ terraced ≈ semi-detached ≈ detached	South East	1	Canterbury
		London	2	Islington
			3	Newham
		East of England	4	Rochford
	Flats/maisonettes ≈ terraced > semi-detached ≈ detached	South East	1	Slough
			2	Fareham
	Flats/maisonettes ≈ terraced	London	1	City of London
Group 3	Flats/maisonettes > terraced ≈ semi-detached > detached	East Midlands	1	Rutland
			2	South Northamptonshire
		East of England	3	Three Rivers
			4	North Norfolk
			5	East Cambridgeshire
			6	Fenland
			7	South Cambridgeshire
			8	Braintree
			9	Chelmsford
			10	Uttlesford
			11	Broxbourne
			12	Dacorum
			13	South Norfolk
			14	Babergh
			15	St Edmundsbury
			16	St Albans
			17	East Hertfordshire
		London	18	Barking and Dagenham

Group No.	Secondary category	Region	No.	LA
			19	Bromley
			20	Ealing
			21	Hounslow
			22	Southwark
			23	Wandsworth
		South East	24	Aylesbury Vale
			25	Wealden
			26	Hart
			27	Winchester
			28	Tonbridge and Malling
			29	Tunbridge Wells
			30	South Oxfordshire
			31	Elmbridge
			32	Guildford
			33	Mole Valley
			34	Reigate and Banstead
			35	Spelthorne
			36	Tandridge
			37	Waverley
			38	Chichester
			39	Horsham
			40	Mid Sussex
		South West	41	Christchurch
			42	East Dorset
			43	North Dorset
			44	West Dorset
			45	Cotswold
			46	Tewkesbury
		West Midlands	47	Herefordshire, County of
			48	Shropshire
			49	Stratford-on-Avon
			50	Warwick
			51	Malvern Hills
		Yorkshire and The Humber	52	Harrogate
			53	Ryedale
	Flats/maisonettes > terraced > semi-detached > detached	East Midlands	1	Harborough
			2	Oadby and Wigston
			3	Daventry
			4	East Northamptonshire
		East of	5	Brentwood

Group No.	**Secondary category**	**Region**	**No.**	**LA**
		England	6	North Hertfordshire
			7	Mid Suffolk
		North East	8	North Tyneside
		South East	9	Milton Keynes
			10	Hastings
			11	Rother
			12	Maidstone
			13	Adur
			14	Worthing
		South West	15	Mid Devon
			16	South Somerset
			17	West Somerset
		West Midlands	18	Solihull
	Flats/maisonettes > terraced ≈ semi-detached ≈ detached	East of England	1	Broadland
		London	2	Hackney
			3	Camden
			4	Hammersmith and Fulham
			5	Kensington and Chelsea
			6	Kingston upon Thames
			7	Richmond upon Thames
			8	Westminster
		South East	9	Brighton and Hove

Appendix C7

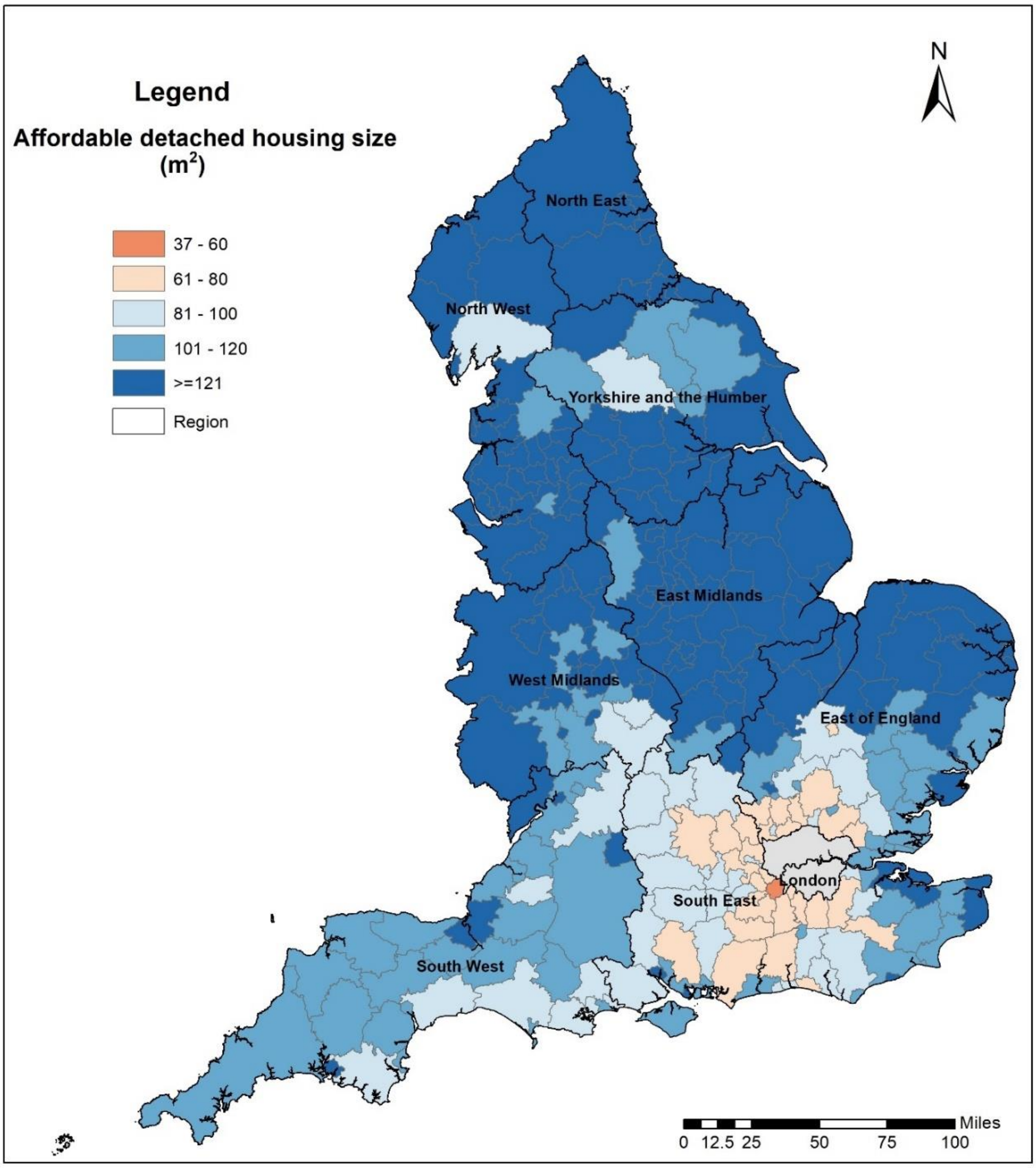


Figure C7.1 The geography of affordable detached property size for typical household HM1 at LA level

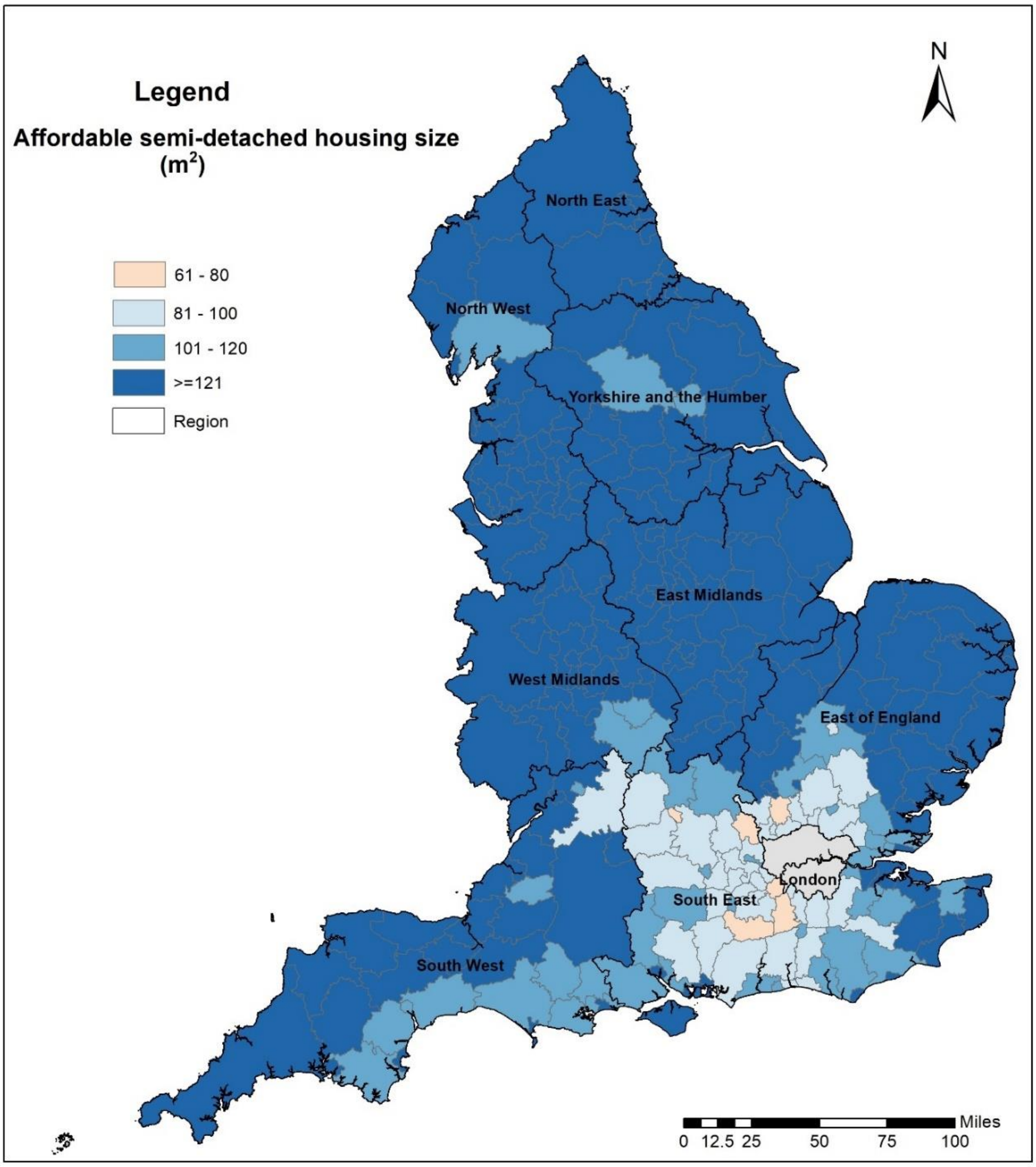


Figure C7.2 The geography of affordable semi-detached property size for typical household HM1 at LA level

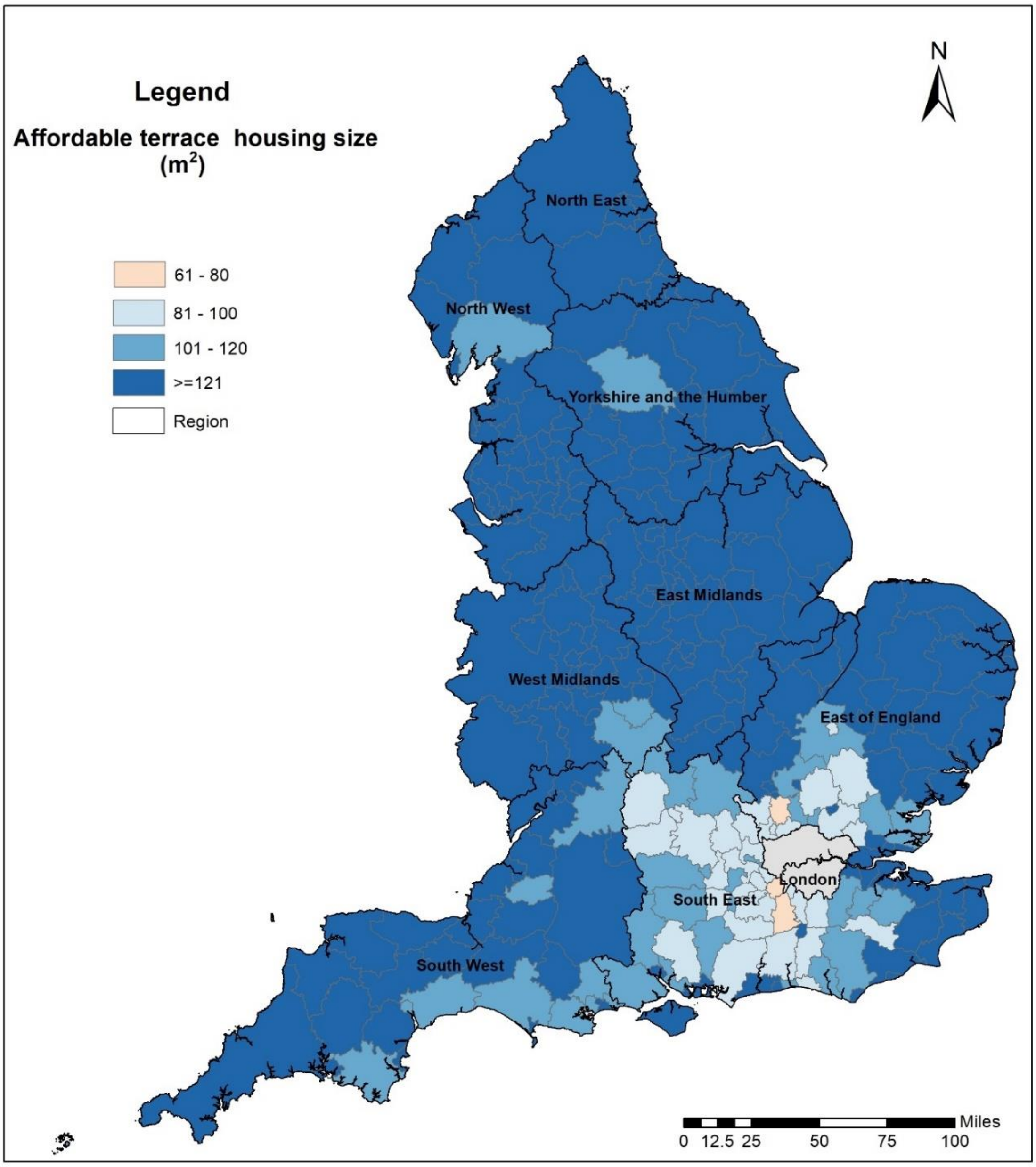


Figure C7.3 The geography of affordable terrace property size for typical household HM1 at LA level

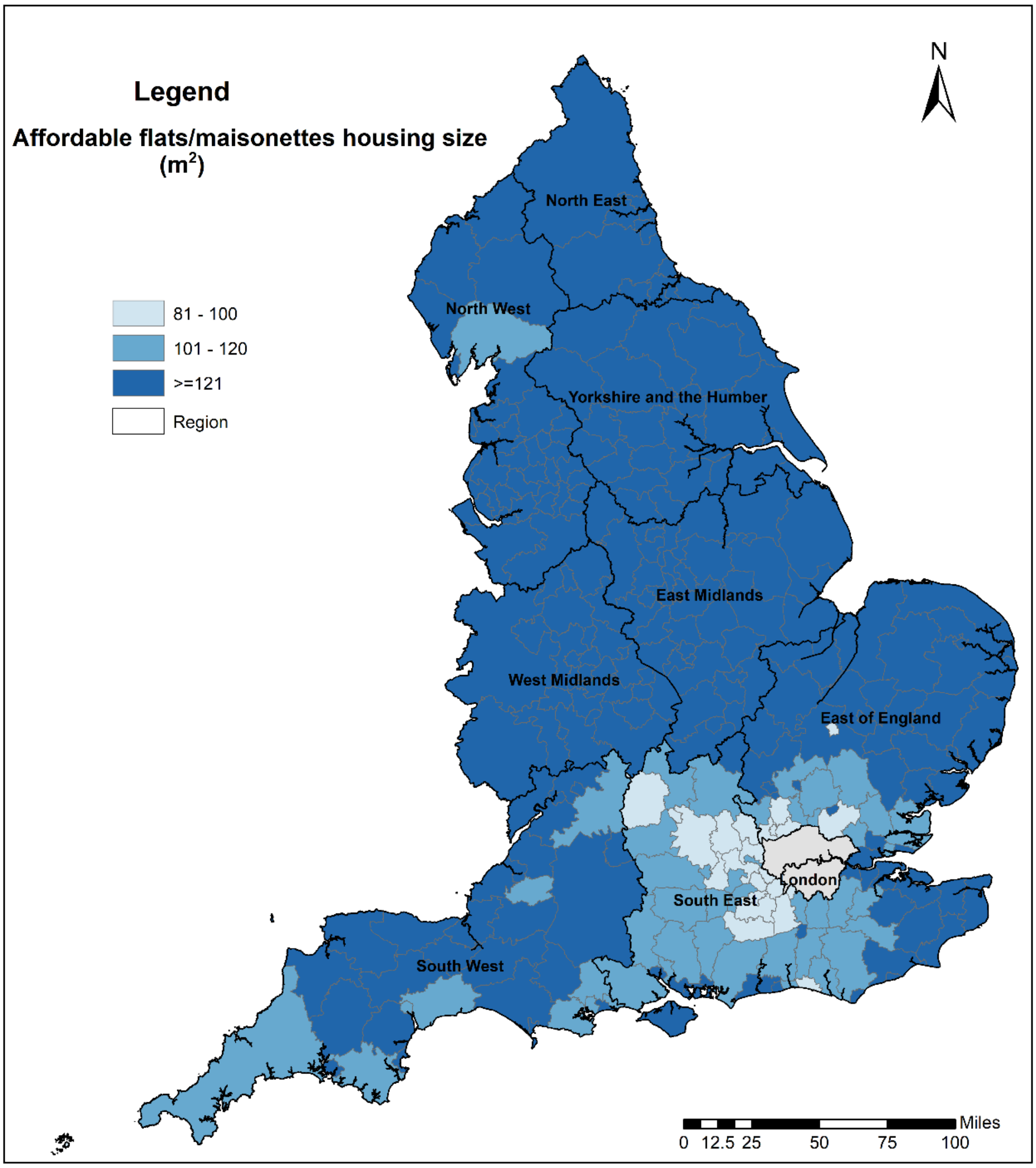


Figure C7.4 The geography of affordable flats/maisonettes property size for typical household HM1 at LA level

Appendix C8

Table C8.1 The change of affordable size by property type for the buyer with £250,000 between 2009 and 2010

LA name	Region	Detached	Semi-detached	Terraced	Flats/ maisonettes
Milton Keynes	South East	-6.65	-8.12	-9.3	-11.24
Waltham Forest	London	-6.01	-6.96	-10.35	-10.2
Barking and Dagenham	London	-4.25	-8	-9.44	-9.76
Corby	East Midlands	-7.83	-10.24	-12.25	-9.57
Lewisham	London	-6.18	-8.34	-8.95	-9.28
Newham	London	-5.76	-7.5	-9.5	-9.19
Croydon	London	-5.33	-7.26	-8.6	-9.02
Harlow	East of England	-4.89	-7.47	-9.03	-8.5
Dartford	South East	-4.56	-7.76	-9.08	-8.48
South Cambridgeshire	East of England	-4.17	-6.16	-5.99	-8.45
Luton	East of England	-6.71	-8.57	-9.45	-8.35
Greenwich	London	-4.27	-7.81	-9.2	-8.26
East Cambridgeshire	East of England	-5.82	-7.65	-7.7	-8.19
East Northamptonshire	East Midlands	-5.56	-8.06	-8.16	-7.84
Harborough	East Midlands	-4.54	-5.45	-6.56	-7.74
Slough	South East	-5.79	-6.37	-7.79	-7.74
Thurrock	East of England	-4.65	-6.25	-8.88	-7.7
Bromley	London	-4.32	-6.34	-7.04	-7.69
Medway	South East	-5.6	-8.27	-10.17	-7.57
Crawley	South East	-5.33	-6.74	-8.23	-7.41
Wellingborough	East Midlands	-6.19	-8.17	-10.45	-7.38
Northampton	East Midlands	-6.65	-7.23	-9.39	-7.36
Bexley	London	-5.68	-7.58	-8.61	-7.3
St Edmundsbury	East of England	-5.02	-6.73	-6.51	-7.24
Haringey	London	-3.69	-4.78	-6.99	-7.2

LA name	Region	Detached	Semi-detached	Terraced	Flats/ maisonettes
Gravesham	South East	-3.97	-6.99	-8.97	-7.17
Sutton	London	-4.58	-6.34	-7.02	-7.02
Enfield	London	-4.56	-6.03	-6.72	-7.01
Coventry	West Midlands	-5	-7.19	-8.14	-6.97
Stevenage	East of England	-5.66	-7.26	-9.37	-6.96
Hackney	London	-4.56	-5.29	-6.88	-6.96
Norwich	East of England	-5.75	-7.2	-7.24	-6.92
Bristol, City of	South West	-6.1	-8.98	-10.13	-6.91
Ealing	London	-4.66	-5.62	-5.94	-6.88
Rugby	West Midlands	-5.02	-7.37	-9.07	-6.81
Hastings	South East	-4.18	-5.8	-6.94	-6.81
Havering	London	-3.73	-6.29	-7.47	-6.79
Lambeth	London	-4.16	-5.73	-6.75	-6.78
Dacorum	East of England	-3.71	-5.91	-6.72	-6.69
Watford	East of England	-5.18	-6.48	-7.43	-6.68
Southwark	London	-4.28	-5.43	-6.55	-6.67
Worthing	South East	-4.07	-5.21	-6.76	-6.66
Forest Heath	East of England	-4.59	-6.42	-9.38	-6.6
Merton	London	-3.26	-5.07	-5.81	-6.56
Maidstone	South East	-4.43	-5.4	-6.15	-6.55
Swale	South East	-5.36	-7.49	-9.17	-6.54
Brent	London	-4.14	-5.54	-6.15	-6.53
Adur	South East	-4.25	-5.01	-6.24	-6.51
Braintree	East of England	-4.03	-6.01	-6.67	-6.49
Ipswich	East of England	-5.61	-8.57	-8.91	-6.46
Dover	South East	-4.65	-6.82	-7.65	-6.44
Thanet	South East	-5.47	-6.73	-7.88	-6.43
Spelthorne	South East	-4.01	-5.37	-5.5	-6.39

LA name	Region	Detached	Semi-detached	Terraced	Flats/ maisonettes
Hounslow	London	-4.07	-5.08	-5.61	-6.36
Broxbourne	East of England	-4.09	-5.73	-6.21	-6.35
Tonbridge and Malling	South East	-3.91	-5.44	-5.58	-6.26
Warwick	West Midlands	-4.2	-5.96	-6.38	-6.22
Southend-on-Sea	East of England	-5.55	-6.85	-8.44	-6.18
Rushmoor	South East	-5.41	-6.19	-7.23	-6.14
Redbridge	London	-4.24	-5.64	-6.14	-6.06
Bedford	East of England	-5.57	-7.39	-7.54	-6.04
Welwyn Hatfield	East of England	-3.26	-6.11	-6.52	-5.99
Hillingdon	London	-4.55	-6.05	-6.41	-5.92
Mid Suffolk	East of England	-3.67	-5.49	-5.73	-5.91
North Hertfordshire	East of England	-4.1	-6.06	-6.92	-5.89
Babergh	East of England	-3.65	-5.09	-6.16	-5.87
Bracknell Forest	South East	-4.52	-5.9	-7.26	-5.86
Oadby and Wigston	East Midlands	-5.3	-6.2	-5.86	-5.84
North East Derbyshire	East Midlands	-2.82	-4.02	-3.09	-5.82
Reigate and Banstead	South East	-3.95	-5.82	-6.08	-5.81
Kingston upon Thames	London	-4.26	-5.03	-5.48	-5.8
South Gloucestershire	South West	-4.83	-6.15	-7.12	-5.79
Brentwood	East of England	-2.54	-4.5	-5.76	-5.78
Central Bedfordshire	East of England	-5.51	-7.38	-7.74	-5.73
Basingstoke and Deane	South East	-4.1	-5.72	-6.75	-5.72
Basildon	East of England	-4.59	-6.34	-8.65	-5.7
Three Rivers	East of England	-3.53	-5.52	-5.44	-5.7
Reading	South East	-4.76	-6.56	-7.54	-5.69
Epsom and Ewell	South East	-4	-5.39	-6.38	-5.68

LA name	Region	Detached	Semi-detached	Terraced	Flats/ maisonettes
Cambridge	East of England	-4.86	-6.1	-5.82	-5.67
Brighton and Hove	South East	-3.3	-4.32	-5.09	-5.63
Surrey Heath	South East	-3.97	-5.42	-5.65	-5.61
Epping Forest	East of England	-3.08	-5	-5.97	-5.58
Harrow	London	-3.66	-5.47	-5.92	-5.56
Manchester	North West	-4.81	-6.54	-6.92	-5.54
East Hertfordshire	East of England	-3.26	-4.6	-5.01	-5.53
Colchester	East of England	-5.11	-6.25	-6.77	-5.5
South Bucks	South East	-3.06	-4.98	-5.34	-5.47
Hertsmere	East of England	-3.56	-5.76	-6.01	-5.45
Tunbridge Wells	South East	-3.34	-4.69	-5.42	-5.45
Portsmouth	South East	-4.26	-6.02	-6.6	-5.39
Daventry	East Midlands	-3.81	-5.02	-7.09	-5.38
St Albans	East of England	-3.02	-4.75	-5.12	-5.37
Gosport	South East	-4.05	-4.39	-5.99	-5.36
Aylesbury Vale	South East	-3.35	-5.43	-6	-5.28
Cherwell	South East	-3.68	-5.67	-5.66	-5.28
Broadland	East of England	-4.3	-5.54	-5.45	-5.21
Barnet	London	-2.87	-4.47	-5.23	-5.16
Wandsworth	London	-2.51	-3.41	-4.18	-5.16
Chelmsford	East of England	-3.73	-5.59	-5.83	-5.14
Mid Sussex	South East	-3.51	-4.67	-5.09	-5.12
Huntingdonshire	East of England	-5.78	-7.07	-7.27	-5.11
Woking	South East	-3.38	-4.88	-5.05	-5.11
Erewash	East Midlands	-5.66	-7.6	-9.01	-5.07
Wycombe	South East	-2.98	-5.29	-5.36	-5.04

LA name	Region	Detached	Semi-detached	Terraced	Flats/ maisonettes
Canterbury	South East	-5.24	-5.53	-5.7	-5.04
Tandridge	South East	-3.4	-5.36	-5.29	-5.01
Swindon	South West	-6.1	-7.51	-8.48	-4.98
Guildford	South East	-2.81	-4.7	-5.08	-4.97
Tower Hamlets	London	-3.57	-5.77	-6.03	-4.97
Fenland	East of England	-6.31	-6.69	-6.33	-4.95
Birmingham	West Midlands	-4.09	-5.4	-5.32	-4.94
Gloucester	South West	-5.63	-6.37	-6.6	-4.9
Tamworth	West Midlands	-4.51	-5.99	-8.16	-4.87
Bath and North East Somerset	South West	-3.55	-4.97	-5.55	-4.84
Uttlesford	East of England	-2.28	-4.27	-4.7	-4.84
Richmond upon Thames	London	-2.52	-3.32	-3.67	-4.8
Ashford	South East	-4.52	-6.11	-6.1	-4.77
Hart	South East	-3.66	-4.31	-5.27	-4.69
Oxford	South East	-3.63	-4.66	-5.07	-4.69
City of London	London	-	-	-4.12	-4.69
Mole Valley	South East	-2.97	-4.54	-4.56	-4.66
Winchester	South East	-2.84	-3.96	-4.49	-4.65
Chiltern	South East	-2.79	-4.39	-4.94	-4.62
Elmbridge	South East	-3.02	-4.29	-4.24	-4.6
North Norfolk	East of England	-3.7	-4	-4.91	-4.59
Redditch	West Midlands	-4.68	-5.26	-5.89	-4.58
Newark and Sherwood	East Midlands	-3.63	-5.16	-5.69	-4.58
Islington	London	-3.08	-4.21	-4.42	-4.51
South Norfolk	East of England	-4.15	-5.17	-5.1	-4.48
Nottingham	East Midlands	-5.84	-8.8	-9.3	-4.48

LA name	Region	Detached	Semi-detached	Terraced	Flats/ maisonettes
Runnymede	South East	-3.58	-4.74	-5.19	-4.44
Windsor and Maidenhead	South East	-3.43	-4.9	-4.77	-4.44
Horsham	South East	-2.64	-4.67	-4.8	-4.41
South Oxfordshire	South East	-2.7	-3.95	-4.33	-4.38
Eastbourne	South East	-4.9	-5.17	-6.31	-4.31
West Berkshire	South East	-3.62	-4.82	-5.35	-4.31
Rushcliffe	East Midlands	-4.63	-5.64	-6.05	-4.29
South Northamptonshire	East Midlands	-4.14	-5.28	-6.09	-4.28
Hammersmith and Fulham	London	-2.8	-3.19	-3.46	-4.26
Lewes	South East	-3.75	-4.63	-4.74	-4.21
Sevenoaks	South East	-3.15	-5.17	-6.19	-4.21
Breckland	East of England	-4.8	-5.79	-7.18	-4.2
Kettering	East Midlands	-6.35	-7.97	-9.56	-4.19
South Kesteven	East Midlands	-5.2	-6.26	-7.37	-4.18
Maldon	East of England	-3.96	-5.78	-5.61	-4.16
Wokingham	South East	-3.89	-5.29	-5.58	-4.16
Broxtowe	East Midlands	-5.63	-6.54	-7.29	-4.12
Shepway	South East	-3.86	-4.68	-5.39	-4.12
North Somerset	South West	-4.75	-5.4	-6.08	-4.1
Rochford	East of England	-4.25	-4.92	-6.14	-4.06
East Hampshire	South East	-3.12	-4.53	-5.37	-4.05
Castle Point	East of England	-5.13	-5.04	-5.87	-4.04
Waverley	South East	-2.57	-3.77	-4.05	-4
Bournemouth	South West	-4.29	-5.55	-5.45	-3.94

LA name	Region	Detached	Semi-detached	Terraced	Flats/ maisonettes
Solihull	West Midlands	-4.19	-5.68	-5.04	-3.93
Trafford	North West	-4.01	-6.46	-7.01	-3.91
Havant	South East	-3.73	-5.26	-6.31	-3.83
South Derbyshire	East Midlands	-4.07	-5.32	-5.88	-3.83
Arun	South East	-3.9	-4.8	-5.26	-3.82
Chichester	South East	-2.38	-3.32	-3.99	-3.78
Nuneaton and Bedworth	West Midlands	-3.91	-5.53	-5.79	-3.72
Peterborough	East of England	-5.33	-6.32	-7.72	-3.69
Southampton	South East	-4.49	-5.47	-5.42	-3.69
Suffolk Coastal	East of England	-3.99	-4.59	-4.53	-3.69
Wealden	South East	-3.11	-4.3	-4.44	-3.68
Tendring	East of England	-4.5	-5.12	-6.32	-3.67
York	Yorkshire and The Humber	-3.12	-4.26	-4.8	-3.67
West Oxfordshire	South East	-2.92	-3.84	-3.79	-3.67
Eastleigh	South East	-3.76	-5.45	-5.12	-3.62
Cheltenham	South West	-4.5	-4.79	-5.4	-3.61
Camden	London	-1.64	-2.35	-3.11	-3.59
Fareham	South East	-3.98	-4.38	-4.91	-3.56
Charnwood	East Midlands	-3.54	-4.7	-4.52	-3.55
Vale of White Horse	South East	-2.95	-4.41	-4.58	-3.54
Waveney	East of England	-4.12	-5.02	-5.31	-3.52
Lincoln	East Midlands	-5.07	-5.94	-8.03	-3.41
Rother	South East	-3.09	-3.77	-4.27	-3.39
Harrogate	Yorkshire and The Humber	-2.42	-3.65	-3.11	-3.39

LA name	Region	Detached	Semi-detached	Terraced	Flats/ maisonettes
King's Lynn and West Norfolk	East of England	-4.43	-5	-5.2	-3.38
Derby	East Midlands	-5.44	-6.52	-5.08	-3.38
Poole	South West	-3.58	-4.75	-4.04	-3.38
Christchurch	South West	-3.17	-4.72	-3.84	-3.33
Leicester	East Midlands	-4.76	-7.7	-7.47	-3.31
Tewkesbury	South West	-3.42	-4.34	-4.91	-3.22
Worcester	West Midlands	-3.37	-3.84	-3.57	-3.14
Test Valley	South East	-2.79	-4.42	-4.89	-3.13
Westminster	London	-1.31	-2.21	-2.06	-3.12
Cotswold	South West	-1.68	-2.92	-2.95	-3.05
Sheffield	Yorkshire and The Humber	-3.67	-4.85	-4.17	-3.04
Stroud	South West	-2.71	-3.96	-3.66	-3
South Hams	South West	-1.33	-2.74	-2.29	-2.99
Wiltshire	South West	-3.74	-4.43	-4.53	-2.98
Rutland	East Midlands	-3.89	-4.79	-3.49	-2.92
Stockport	North West	-4.4	-5.63	-5.81	-2.77
East Dorset	South West	-3.03	-3.74	-4.25	-2.75
Exeter	South West	-3.19	-3.75	-3.45	-2.75
Isles of Scilly	South West	-2.76	-3.57	-2.92	-2.75
Gedling	East Midlands	-4.71	-6.5	-5.94	-2.73
Mendip	South West	-3.68	-4.02	-4.73	-2.71
Kensington and Chelsea	London	-1.74	-1.97	-1.42	-2.68
Lichfield	West Midlands	-3.42	-3.91	-4.84	-2.68
Stratford-on-Avon	West Midlands	-2.83	-3.8	-3.92	-2.62
Purbeck	South	-2.26	-3.9	-3.56	-2.5

LA name	Region	Detached	Semi-detached	Terraced	Flats/ maisonettes
	West				
Leeds	Yorkshire and The Humber	-3.1	-3.86	-1.56	-2.43
South Somerset	South West	-1.85	-2.97	-2.68	-2.39
Mid Devon	South West	-2.36	-2.28	-1.97	-2.36
North West Leicestershire	East Midlands	-4.01	-5	-5.72	-2.25
New Forest	South East	-2.87	-4.25	-4.61	-2.25
Malvern Hills	West Midlands	-2.75	-3.17	-3.2	-2.22
Torridge	South West	-1.68	-2.37	-2.46	-2.12
Salford	North West	-3.5	-5.87	-9.91	-2.12
Cannock Chase	West Midlands	-4.4	-5.29	-4.95	-2.12
Teignbridge	South West	-2.11	-2.75	-2.6	-1.97
Bolsover	East Midlands	-4.02	-5.72	-5.79	-1.96
North Dorset	South West	-2.24	-3.11	-2.82	-1.96
North Warwickshire	West Midlands	-2.6	-4.52	-3.66	-1.96
Blaby	East Midlands	-4.47	-5.31	-6.36	-1.93
Plymouth	South West	-3.07	-4.13	-4.21	-1.91
West Lindsey	East Midlands	-3.57	-4.23	-0.56	-1.6
Warrington	North West	-3.22	-4.93	-5.02	-1.48
Sedgemoor	South West	-2.57	-3.23	-4.59	-1.44
Hinckley and Bosworth	East Midlands	-4.09	-4.59	-5.2	-1.43
East Devon	South West	-2.07	-2.68	-2.7	-1.41
West Dorset	South West	-2.47	-2.77	-2.54	-1.38
East Staffordshire	West Midlands	-3.68	-5.03	-5.19	-1.37
Great Yarmouth	East of England	-4.73	-5.02	-4.42	-1.36
Cheshire East	North	-3.14	-3.97	-4.59	-1.33

LA name	Region	Detached	Semi-detached	Terraced	Flats/ maisonettes
	West				
Torbay	South West	-2.58	-2.75	-3.58	-1.24
Weymouth and Portland	South West	-2.79	-3.18	-3.3	-1.17
Melton	East Midlands	-3.34	-4.99	-5.5	-1.14
Chesterfield	East Midlands	-3.99	-5.96	-5.76	-1.12
Wolverhampton	West Midlands	-2.37	-4.69	-4.25	-1.07
Isle of Wight	South East	-1.48	-2.4	-2.36	-1.02
North Tyneside	North East	-2.77	-4.19	-3.17	-0.96
Cheshire West and Chester	North West	-2.6	-3.58	-3.86	-0.94
Boston	East Midlands	-4.6	-5.75	-6.69	-0.93
South Staffordshire	West Midlands	-1.8	-2.97	-3.52	-0.92
Wychavon	West Midlands	-1.94	-2.67	-3.07	-0.9
Kirklees	Yorkshire and The Humber	-1.88	-2.54	-2.43	-0.89
Shropshire	West Midlands	-1.56	-2.29	-1.76	-0.86
West Devon	South West	-1.85	-2.94	-2	-0.85
Amber Valley	East Midlands	-4.13	-5.96	-6.4	-0.85
North Kesteven	East Midlands	-4.87	-5.09	-5.47	-0.81
Derbyshire Dales	East Midlands	-2.85	-3.63	-2.96	-0.76
High Peak	East Midlands	-2.88	-4.38	-3.49	-0.7
Stafford	West Midlands	-2.96	-3.5	-3.07	-0.62
Bromsgrove	West Midlands	-2.79	-4.28	-3.68	-0.61
Taunton Deane	South West	-2.66	-3.29	-2.55	-0.57
Hambleton	Yorkshire and The Humber	-1.69	-1.88	-1.63	-0.56
Bury	North West	-2.84	-4.59	-4.77	-0.55

LA name	Region	Detached	Semi-detached	Terraced	Flats/ maisonettes
Staffordshire Moorlands	West Midlands	-2.98	-4.29	-2.83	-0.52
Wyre Forest	West Midlands	-2.17	-2.69	-3.15	-0.46
Cornwall	South West	-1.57	-2.07	-2.27	-0.44
Newcastle upon Tyne	North East	-2.25	-3.53	-1.92	-0.43
Richmondshire	Yorkshire and The Humber	-1.17	-1.75	-1.32	-0.42
Rotherham	Yorkshire and The Humber	-3.6	-4.86	-1.64	-0.31
Mansfield	East Midlands	-4.59	-6.81	-6.95	-0.29
Newcastle-under-Lyme	West Midlands	-3.76	-5.14	-4.69	-0.26
Ashfield	East Midlands	-5.49	-8.23	-8.64	-0.23
North Devon	South West	-2.22	-2.39	-3.11	-0.22
Sandwell	West Midlands	-2.99	-5.12	-5.19	-0.21
Tameside	North West	-2.85	-4.23	-4.34	-0.15
West Somerset	South West	-1.92	-2.85	-2.18	-0.09
Craven	Yorkshire and The Humber	-1.3	-2.03	-2.3	-0.08
Halton	North West	-2.84	-3.4	-3.13	-0.04
South Holland	East Midlands	-6.73	-7.18	-8.19	0.01
Liverpool	North West	-2.1	-4.19	-2.22	0.12
Selby	Yorkshire and The Humber	-3.09	-3.49	-4.45	0.15
Walsall	West Midlands	-3.64	-4.75	-4.21	0.23
South Lakeland	North West	-0.27	-0.64	-0.97	0.23
Herefordshire, County of	West Midlands	-2.27	-3.54	-3.16	0.27
Lancaster	North West	-0.63	-2.12	-0.97	0.45
Ryedale	Yorkshire	-0.98	-1.61	0.13	0.45

LA name	Region	Detached	Semi-detached	Terraced	Flats/ maisonettes
	and The Humber				
Eden	North West	0.46	0.2	0.95	0.49
Ribble Valley	North West	0.5	-0.9	-1.36	0.52
Wirral	North West	-2.51	-3.54	-2.2	0.6
South Ribble	North West	-1.21	-1.14	-1.82	0.62
Dudley	West Midlands	-2.86	-3.57	-3.42	0.74
Fylde	North West	-1.2	-1.75	-3.4	0.77
Stoke-on-Trent	West Midlands	-4.43	-5.94	-5.54	0.8
Gateshead	North East	-1.89	-2.63	-1.63	0.83
Calderdale	Yorkshire and The Humber	-2.22	-2.17	-0.31	0.96
Rossendale	North West	-0.84	-2.29	0.61	1.07
Doncaster	Yorkshire and The Humber	-3.17	-3.68	0.27	1.22
East Lindsey	East Midlands	-3.1	-2.75	-4.15	1.35
Scarborough	Yorkshire and The Humber	-0.31	-0.69	1.33	1.41
Sefton	North West	-2.16	-3.04	-1.35	1.42
Forest of Dean	South West	-2.72	-3.39	-3.57	1.45
Rochdale	North West	-1.08	-1.75	-0.18	1.47
Bradford	Yorkshire and The Humber	-1.47	-1.18	2.55	1.6
East Riding of Yorkshire	Yorkshire and The Humber	-2	-2.41	-2.14	1.8
Oldham	North West	-1.87	-2.36	-1.07	1.82
Kingston upon Hull, City of	Yorkshire and The Humber	-3.82	-4.68	-4.5	1.9
Bolton	North	-2.38	-2.68	-3.04	2

LA name	Region	Detached	Semi-detached	Terraced	Flats/ maisonettes
	West				
Preston	North West	-1.81	-1.68	-0.49	2.09
West Lancashire	North West	-1.39	-2.04	-1.92	2.19
St. Helens	North West	-1.6	-1.92	-0.05	2.54
North East Lincolnshire	Yorkshire and The Humber	-1.88	-2.86	-0.89	2.59
Burnley	North West	-1.4	-0.46	3.67	2.75
Wakefield	Yorkshire and The Humber	-1.92	-2.39	-1.91	2.82
Stockton-on-Tees	North East	-2.09	-1.84	0.11	2.89
Redcar and Cleveland	North East	-0.93	-0.53	2.69	2.92
Allerdale	North West	0.44	-0.64	1.85	2.97
Wyre	North West	-0.35	-0.34	0.29	2.98
Darlington	North East	-2.21	-1.87	1.4	2.99
Wigan	North West	-2	-2.66	-3.66	3.03
Bassetlaw	East Midlands	-3.54	-4.94	-4.09	3.07
Northumberland	North East	-0.87	-1.53	0.06	3.18
North Lincolnshire	Yorkshire and The Humber	-2.39	-2.3	-2.47	3.19
Telford and Wrekin	West Midlands	-3.13	-3.1	-2.65	3.23
South Tyneside	North East	-1.12	-2.55	-1.1	3.24
Carlisle	North West	0.03	-1.37	0.17	3.24
Barnsley	Yorkshire and The Humber	-2.51	-2.7	0.85	3.25
Blackburn with Darwen	North West	-1.16	-0.53	1.66	3.26
Knowsley	North West	-1.78	-2.39	-3.03	3.29
Pendle	North West	-1.69	-2.37	0	3.41

LA name	Region	Detached	Semi-detached	Terraced	Flats/ maisonettes
Copeland	North West	-2.2	-2.5	-1.43	3.69
Chorley	North West	-1.33	-1.26	-1.62	4.22
Blackpool	North West	0.25	-0.87	1.39	4.5
Middlesbrough	North East	-1.76	-0.65	2.73	4.53
Hyndburn	North West	-1.32	0.12	0.76	4.72
County Durham	North East	-1.12	-0.23	4.5	6.65
Barrow-in-Furness	North West	-1.56	-2.59	-0.79	6.97
Sunderland	North East	-1.58	-1.81	1.6	7.13
Hartlepool	North East	-1.1	-0.83	7.23	7.5

Table C8.2 A summary of LA's annual affordable property size change associated with standard bedrooms

Item	Detached	Semi-detached	Terraced	Flats/maisonettes
Lost a single bedroom	Corby	Corby	Wellingborough	Milton Keynes
		Bristol, City of	Waltham Forest	Waltham Forest
		Nottingham	Medway	Barking and Dagenham
		Luton	Bristol, City of	Corby
		Ipswich	Salford	Lewisham
		Lewisham	Kettering	Newham
		Medway	Newham	Croydon
		Ashfield	Luton	Harlow
		Wellingborough	Barking and Dagenham	Dartford
		Milton Keynes	Northampton	South Cambridgeshire
		East Northamptonshire	Forest Heath	Luton
		Barking and Dagenham	Stevenage	Greenwich
		Kettering	Milton Keynes	East Cambridgeshire
		Greenwich	Nottingham	East Northamptonshire
		Dartford	Greenwich	Harborough

Item	Detached	Semi-detached	Terraced	Flats/maisonettes
		Leicester	Swale	Slough
		East Cambridgeshire	Dartford	Thurrock
		Erewash	Rugby	Bromley
		Bexley	Harlow	Medway
		Swindon	Erewash	
		Newham	Gravesham	
			Lewisham	
			Ipswich	
			Thurrock	
			Basildon	
			Ashfield	
			Bexley	
			Croydon	
			Swindon	
			Southend-on-Sea	
			Crawley	
			South Holland	
			Tamworth	
			East Northamptonshire	
			Coventry	
			Lincoln	
			Thanet	
			Slough	
			Central Bedfordshire	
			Peterborough	
			East Cambridgeshire	
			Dover	
			Bedford	
			Reading	
Lost a double bedroom	-	-	Corby	-
Gain one more single bed room	-	-	-	Hartlepool