Newcastle University e-prints

Date deposited: 15th April 2013

Version of file: Author final

Peer Review Status: Unknown

Citation for item:

Jones C, Coombes M, Dunse N, Watkins D, Wymer C. [Tiered housing markets and their relationship to labour market areas](#). *Urban Studies* 2012, **49**(12), 2633-2650.

Further information on publisher website:

<http://www.sagepub.com>

Publisher's copyright statement:

© 2012, Pion Ltd.

[Authors] may post on any non-commercial* repository or website* the version of your article that was accepted for publication – ‘**version 2**’. The article may not be made available earlier than 12 months after publication in the Journal issue and may not incorporate the changes made by SAGE after acceptance.’

The definitive version of this article is available at:

<http://dx.doi.org/10.1177/0042098011435847>

Always use the definitive version when citing.

Use Policy:

The full-text may be used and/or reproduced and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not for profit purposes provided that:

- A full bibliographic reference is made to the original source
- A link is made to the metadata record in Newcastle E-prints
- The full text is not changed in any way.

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

**Robinson Library, University of Newcastle upon Tyne, Newcastle upon Tyne.
NE1 7RU. Tel. 0191 222 6000**

Tiered Housing Markets and their Relationship to Labour Market Areas

**Colin Jones¹, Mike Coombes², Neil Dunse¹,
David Watkins¹ and Colin Wymer²**

¹ School of Built Environment, Heriot-Watt University

² Centre for Urban and Regional Development Studies, Newcastle University

Tiered Housing Markets and their Relationship to Labour Market Areas

Abstract

This paper develops a tiered geography of local housing market areas (HMAs) and examines their relationship with labour market areas. The argument is based on the theoretical understanding of the basis of HMAs. The relationships between the tiers and labour market areas are explored. Drawing on this understanding, the empirical research generates sets of different potential geographies of HMAs for England based on an algorithm that applies a range of criteria linked to the degree of closure of migration and/or commuting patterns. A range of theoretically appropriate criteria then enable the different geographies to be assessed. In particular, standardised house prices in neighbouring HMAs are tested to assess whether they are different using hedonic multiple regressions and a Chow test is used to see if they generate statistically different coefficients. Finally conclusions are drawn on the validity of the approach developed. The empirical work is on England and the datasets drawn from the Population Census and the Land Registry.

1. Introduction

This paper develops a tiered geography of local housing market areas (HMAs) and examines their relationship with local labour market areas. The justification for this task lies partly in the growing use of HMAs as a planning tool in the UK and other countries, but also because using a valid HMA geography is an essential prerequisite for a full understanding of local housing market dynamics. Underlying this issue is the modifiable areal unit problem (Openshaw and Taylor 1981) which determines that the results of spatial analyses will depend at least in part on which areas are used for those analyses. As recognised by Briant *et al* (2010), the appropriate response is to use areas that are robustly defined with respect to the theoretical basis of the relevant concepts. In the USA the usual response has been to use metropolitan areas whose definitions are rooted in labour market areas. This begs the question of whether housing and labour market areas are necessarily the same, so this paper tackles that question as part of a broader attempt to meet the challenge of creating robust HMA definitions based on the concept of the sub-regional housing market.

The definition of HMAs has often been ignored in empirical studies of local housing markets, following the pioneering analyses of the spatial structure of urban housing sub-markets by Straszheim (1975) and Schnare and Struyk (1976). The subsequent burgeoning literature was reviewed recently by Jones and Watkins (2009), revealing that although over time there have been refinements such as improved stratification procedures and greater data availability, most studies simply adopt administrative boundaries as a 'given' for the overall HMA. Similar issues apply to the measurement of local supply elasticities, as with the use of local authority boundaries by Green *et al* (2005) and Pryce (1999). It is clear that the results will be very dependent on the area to which the analysis is applied: for example, a supply elasticity measured on an urban core that excludes its suburbs will be prone to give a biased estimate. Gyourko (2009) in his review of housing supply concludes that an important area for future research is to calculate robust local market supply elasticities, and the argument here is that this in turn requires that such estimates are based on genuine housing market areas.

There have been a small number of academic studies that have considered the definition of HMAs but the motivation for this paper requires a more fundamental review than yet exists in the literature. Links between housing market behaviour and the labour market have tended to focus on residential mobility and job mobility, for example Kan (2002) and Pinto (2002). In particular, it is necessary to revisit conceptual issues and to focus on the links between HMAs and labour market areas. The paper seeks to explore the relationships, both theoretical and empirical, between these two basic functional economic geographies (Fox and Kumar 1965), labour and housing market areas. In doing so it sets out a practical and consistent national geography of HMAs for England, after first showing that there are no easy answers to the construction of such a geography in the face of both theoretical and practical challenges.

The paper begins by comparing previous studies of local labour and housing market areas, noting the dramatic difference in the volume of the two sets of previous academic studies. The following section provides an explanation of the theoretical

perspective on the housing market that underpins the case for a tiered view of HMAs. Next the empirical research generates sets of different potential geographies of HMAs for England based on an algorithm that groups areas and applies a range of criteria derived from the theoretical underpinnings. The next section tests whether there are statistical differences in the results of standardised house price tests between contiguous pairs of individual HMAs in the different geographies. The empirical work draws on data from the Population Census and the Land Registry. Finally some general and specific conclusions are outlined.

2. Defining Labour and Housing Market Areas

Housing and labour markets are unusual in that it is the consumer who moves not the product. Studies of the geography of local labour markets are of longer standing than those of HMAs. In the USA there has been a longstanding recognition of the need to define metropolitan areas in terms of labour market areas, with this approach more recently extended to provide a diverse set of functional economic areas covering the whole territory (see for example Dahmann (1999) on the question of extending labour market definitions beyond metropolitan areas). Labour market areas are defined by the analysis of commuting patterns, given the necessary data. Coombes *et al* (1986) and Tolbert and Killian (1987) exemplify these analyses in the UK and USA respectively, and the UK approach – which defines Travel-to-Work Areas (TTWAs) – has subsequently been used in several countries across the world, such as Spain (Casado-Diaz 2000). The logic of analysing commuting patterns is that these are the spatial expression of Cournot's underlying principle of spatial arbitrage as the process that creates a market. Thus the TTWA is a labour market area because within it the buyers and sellers of labour interact to establish wage rates (prices). In practice such an area will never be completely closed, especially in more heavily urbanised areas like much of the England where people can readily live and work in different cities. As a result, it is necessary to select a level of commuting closure to define sub-regional labour market areas: there is no theoretical basis for choosing one level or another. For example, the level applied to define TTWAs has varied through the years; it is currently 66.7% (leading to 140 TTWAs in England). The parallels between housing and labour markets as sub-regional functional areas suggest that many of the defining features of labour markets as spatial markets may well also apply to defining HMAs.

There are only a few published academic studies that have identified HMA geographies and discussed the relationship with labour market areas. All these studies have been in the UK, with one of the first being an application of the TTWA method of analysis to migration data (Coombes 2000, p.1510-1511). The rationale for analysing migration flows to define HMAs is that the guiding principle of spatial arbitrage implies that the geography of HMAs will be formed through the pattern of movement between where people move to and where they originate from. The first full academic study taking this approach was Jones (2002): boundaries of HMAs were defined by analysing the migration patterns of home buyers. The spatial focus of the analysis is the area broadly defined as mainland west central Scotland centering on Glasgow. The migration data is derived from the Land Registry covering the ten year period 1984 to 1993. The approach is based on grouping **settlements** to establish HMAs through the analysis of migration interaction. These

settlements range in size from the city of Glasgow to small villages, their attraction is their internal coherence. The HMA is based on the notion that it can be defined as a contiguous area comprising a settlement or group of settlements with a high degree of housing market self-containment, and where in-migration from outside the immediate HMA is of only minor significance.

The grouping of settlements is undertaken using an iterative algorithm containment benchmark: at least 50% internal migration, or in-migration from an adjacent HMA equivalent to less than 5% of the destination market. These criteria yielded a set of HMAs which were seen to have an embedded relationship with TTWAs, and it was this link between housing and labour market area definitions that led Jones (2002) to select these particular criteria after showing how varying the criteria leads to the geography of HMAs changing substantially..

A delineation of HMAs for the North West of England encompassing Manchester and Liverpool was undertaken by Brown and Hincks (2008). Although these HMA definitions were based primarily on migration data, part of the process involved consulting estate agents. Comparison of the HMAs derived on this basis with the TTWAs in the region revealed similarities in most urban areas. There were greater differences between HMA and TTWA boundaries in more rural areas, indicating that the relationship between housing and labour markets may vary by type of area.

Coombes (2009) directly tested the relationship between housing and labour markets by analysing the Census migration data with the method used to define TTWAs. The fact that the results varied markedly in different parts of the country prompted the innovation of taking account of non-movers, but the conclusion was that the results were unsatisfactory, even after many variations to the closure levels (ie. the proportion of migrants 'allowed' to cross boundaries of HMAs). .

All these studies apply a migration closure approach, although with different algorithms and datasets. Various closure criteria were applied, partly but not entirely because of the different data sources. The studies highlight there is no a priori theory to guide the choice of the closure criteria for migration flows. None of the results from studies suggest that housing market areas come close to matching "1 to 1" with labour market areas across a mix of rural and more urbanised areas, although there was a clear expectation that this could occur when policy guidance was issued for the definition of HMAs in England (Communities and Local Government, 2007).

The conclusions from this brief review are threefold:

- Unlike the definition of labour market areas, the most appropriate way of defining HMAs is under-researched: indeed their definitions are scarcely an issue for academic debate outside Britain, despite the fact that the areas used will affect the results of any sub-regional analysis of housing finance.

- The second conclusion is that there does not yet exist an established approach to defining HMAs; moreover the methods which have been tried have shown that there is no 'natural' level to set for the key value of closure of migration flows (ie. the proportion of flows that must not cross the boundary of the HMA if it is to be deemed a sufficiently separate sub-regional market).
- The third conclusion is that the existing empirical analyses in Britain do not support the idea that labour and housing market areas are effectively substitutable for each other: although it is *possible* to adjust the closure criteria to create similar HMA and labour market area definitions in one type of area (eg. metropolitan regions), doing so will lead the two sets of areas to be very different for other parts of the country (eg. rural areas).

From this starting point, the necessary step is to return to the concept of housing markets and then to develop methods for the empirical definition of HMAs which are rooted in this theoretical understanding.

3. Theoretical Perspective

This section argues that the system of local housing markets can be seen as series of tiers. It begins by reviewing the theory of urban housing markets that centres on the role of journey to work as a key influence. It then focuses on the role of spatial arbitrage in moulding the nature of housing markets via household migration. When households, whether they have a member who is working or not, move the process of price bidding is not only the internal housing market dynamic but it is argued also the basis for determining the boundaries of HMAs.

The essentials of the theory of urban housing markets were developed by Alonso (1964), Muth (1969) and Evans (1973). They develop the concept within an urban area that is characterised by the following key assumptions:

- the town or city occupies a featureless plain, so any topographical features that might distort key relationships are ignored,
- employment is concentrated in the city centre, the central business district, and households make a fixed number of work trips a week.

The housing market in this model is assumed to have perfect information and that households then make bids for particular locations and through this process a price surface emerges. In this housing market the law of one price holds but prices vary with distance or accessibility from the city centre because in deciding the price to bid households take into account the transport cost from any location to the central business district.

Households are prepared to bid a higher price for an equivalent house (of the same size etc) in more accessible locations with lower travel costs than one in locations further out. This basic model assumes that all housing quality (including types) is the same and that there are no neighbourhood preferences within an urban area. Within the model, known as the 'access-space' model, the equilibrium price of housing per square metre declines with distance from the city centre. Muth (1969) demonstrates

mathematically within the confines of the strict assumptions that for a stable long run equilibrium the house price gradient has to be a negative exponential function with house prices decreasing at a slower rate with distance from the city centre.

The model presumes a dominant city or town centre that represents the key point of accessibility and the major locus of urban employment. The urban housing and labour markets are the same. The current pattern of settlements and commuting does not conform to these assumptions. First, the urban systems of modern countries do not comprise a series of independent towns with separate commuting patterns. In addition, within cities commuting trips are no longer necessarily only from suburbs to city centre because sub-centres increasingly exist within most city regions (McMillen and Smith 2003). Outside city regions there are often polycentric sub-regions with several towns where the key accessibility relationship is linked not to the centre of the town with the largest population but the point of greatest 'regional' accessibility within the inter-urban road network.

Notwithstanding these differences between the hypothetical and actual urban system and its commuting patterns, empirical price studies consistently find a significant distance decay function from central urban locations (see Jones *et al* 2009, for example). This finding implies that the essential dynamic of the access-space model holds under less restrictive conditions, and the journey to work is the key force in shaping local spatial housing markets. The corollary is that the limits to local HMAs are determined by travel to work patterns. In other words the outer boundaries of HMAs are determined by the distances travelled by the longest commuters in different directions from a dominant accessibility point. Within this perspective spatial house price arbitrage occurs as households move within this wider labour market area. Thus the labour market area of longer-distance commuters sets the outer bounds for housing market areas, which are here called *Framework HMAs*.

There are key qualifications to these conclusions. First, the access-space model represents a long term equilibrium view of the housing market so HMAs defined by commuting patterns is best viewed as a framework within which spatial housing market processes operate. Second, the simplifying assumptions of the access-space model neglect important dimensions of the housing market and its short term dynamics, namely that households have preferences for different house types and neighbourhoods and areas, and that the housing stock is differentiated in terms of housing quality and types and (relatively) fixed at any particular location. Finally the assumption of a unitary housing market within an urban area in which the law of one price holds has also been the subject of considerable academic debate and challenge through the literature on submarkets beginning with Straszheim (1975) and Schnare and Struyk (1976). Many factors linked to restricted household mobility and (in numerous regions including most of Britain) the slow response of new house building to price rises. The result is that price differences across different parts of an urban market which the model assumes will be short term in practice persist into the long term. In other words the extent of spatial arbitrage in the *Framework HMA* that was defined by commuting is fragmented, implying that several HMAs can co-exist within that wider area.

The heterogeneity of housing, the diversity of neighbourhoods and locations within a sub-region and the short distances often moved by households can thus produce subsystems or layers within a *Framework HMA* due to the differences not being arbitrated away. There have been different approaches to the measurement of these subsystems that are defined by constraints on spatial arbitrage. The first approach analyses migration patterns between and within settlements: if an area has a degree of self-containment in the migration flows, then the fluidity of spatial arbitrage within that area will persist alongside a quasi-independence from other parts of the *Framework HMA*. This is the approach to defining HMAs taken in the British studies outlined above. The second approach considers the outcomes of this quasi-independence, so that the lack of spatial arbitrage should result in differences in the prices of a standardised house in each subsystem. This is tested by using hedonic price analysis and is the basis of many submarket studies.

To illustrate this tiered nature of the housing market, one starting point is to consider household movement through the family life stages and the range of substitutes and locations households consider when moving home. City centre living, usually in a flat, has become popular for childless households in their twenties and thirties. Later in life households with children often prefer a home with the use of a garden, or place greater emphasis on neighbourhood factors such as school catchment areas (assuming the work search areas remain unchanged). The price a household is prepared to pay for a specific house will reflect a combination of its structural characteristics and the neighbourhood in which it is located. Although this price will in the long-term be determined by reference to the wider fundamental spatial house price structure of the whole *Framework HMA*, the spatial arbitrage processes are limited by actual migration patterns, leading to the possibility of defining a separable set of smaller areas that are here termed *Local HMAs*.

This spatial perspective can be further disaggregated to recognise neighbourhood or house type *submarkets*. The concept of the submarket implies that the urban housing market may be segmented on both the demand and supply side of the market. From a demand perspective households may form distinct 'consumer groups' with associated housing preferences and tastes that are in turn linked to stage in the family life cycle, size and composition, and socio-economic status. These 'consumer groups' may also have similar constraints in their search and information costs. In parallel the housing stock (supply) is also segmented into product groups (MacLennan *et al*, 1987) that represent relatively homogenous dwellings and hence close substitutes to the demanders of housing. The existence of submarkets implies segmented demand is matched to the differentiated housing stock and results in differential prices to be paid for given attributes in different market segments. In this way premiums for a particular neighbourhood and/or house type are derived.

To summarise: the constraints on market adjustment or spatial arbitrage between *Local HMAs* (and even *submarkets* where these exist) means that standardised house prices in different parts of the same *Framework HMA* can be very different. Spatial arbitrage occurs, but indirectly and with a time lag. Excess demand for particular dwellings (and their close substitutes) will drive prices in that *Local HMA* upward,

but may not affect other *Local HMAs*. The result is that different parts of a *Framework HMA* may have very different house price structures, and hence different house price inflation trends and levels of affordability. This also means building new houses in one part of *Framework HMA* may not necessarily address an affordability problem due to supply shortages in a particular *Local HMA* if it does not lead to a redrawing of migration patterns. To achieve this will require a sensitive approach to the location of such new housing taking into account transport networks for example and demands a focus on *Local HMAs* embedded within their *Framework HMA*.

The discussion has therefore argued that there are three potential tiers to the structure of HMAs.

1. *Framework HMA* defined by long distance commuting flows
2. *Local HMAs* defined by migration patterns
3. *Housing Submarkets* defined in terms of neighbourhood and/or house type price premiums

This theoretical analysis creates the guidelines for our approach to identifying *Framework HMAs* and *Local HMAs*: to restrict the scope of this paper to some degree, we do not aim to define *Housing Submarkets* as well. In particular, the *Framework HMAs* definitions should be based on the analysis of commuting, whilst the definitions of *Local HMAs* will derive from migration patterns. It is possible that these two sets of areas may partially collapse into a single set of boundaries, or may not closely align with each other where the relationship between migration and commuting 'on the ground' is complex. It is most likely that *Framework HMAs* will be considerably larger than *Local HMAs* where long distance commuting is widespread (eg. around major conurbations). By contrast, *Local HMAs* could actually be larger than *Framework HMAs* in some rural areas where many of the migrants are retired and so not part of the local labour market, and where commuting patterns for most workers are localised.

4. Research Method

The empirical research constructs a tiered HMA geography in stages by applying a grouping algorithm to sets of commuting and migration flows. The final tiered geography is derived by constraining or embedding the boundaries of *Local HMAs* within those of *Framework HMAs*. As was mentioned earlier, a key issue is that there is no a priori basis for the degree of closure of commuting and migration which will be required of the tiers of HMAs. In fact even if a whole country was considered a single HMA it would not have 100% closure of either its commuting or its migration flows. It should also be noted that the containment levels of the two flows are not directly comparable because commuting is a daily activity whereas migration is periodic. The analysis therefore develops a strategy that seeks to address this issue that ensures the outcomes arise from nationally consistent criteria and are theoretically sound.

The choices of closure levels are based on two criteria. First, the choice can be guided by the theory above which argues that in more urban areas it is likely that *Local*

HMA_s will be embedded within *Framework HMA_s*. To achieve this, the analyses examine the inter-relationship between the spatial patterns produced using different potential levels for the closure of the commuting/migration flows analysed. Second, selected geographies are then compared by testing how many contiguous pairs of individual HMA_s have statistical differences in standardised house prices. In particular standardised house prices in neighbouring HMA_s are tested to assess whether they are different using hedonic multiple regressions and a Chow test to see if they generate statistically different coefficients. This procedure follows the standard submarket testing first developed by Schnare and Struyk (1976) and also permits a comparison of the reduction in regression variance achieved by the HMA regressions compared with the national equation.

The grouping algorithm applied uses the TTWA definition method (Coombes 2010) which does not impose any structure (eg. core-periphery) but simply identifies all clusters of flows of any form. In this case the algorithm groups commuting or migration flows between wards identified in the 2001 Census. The algorithm seeks to identify as many as possible separate areas which meet the key criterion of the set level of closure (ie. the proportion of the flows analysed which both start and end within the same area). It does this by grouping areas in whatever way minimises the number of flows that cross them. Different levels of migration or commuting closure are considered.

Migration data in Britain tends to be strongly dominated by the numerous lengthy moves of students, who are not directly relevant to this research. The published Census migration data does not cross-tabulate households by either small age groups or whether the person was a student, so the effect of students on the research is reduced by the use of a customised dataset of Moving Group Reference Persons (MGRPs) specifically provided by the Office of National Statistics. This dataset records the tenure of migrants and excluded all people aged under 25. It should be noted here that the definition of MGRPs covers many people who are not heads of households: for example, a 25 year old returning to the parental home will be a single person moving group, and if the parental home is owner-occupied then this 25 year old will be recorded as an owner-occupying MGRP because the same tenure characteristic applies to all household members. The research presented is based on the dataset covering all 25(+) MGRPs but the results for the owner occupied sector only are similar in fact (due not only to owner-occupiers being the majority of all migrants, but also because the average length of their moves lies midway between the short distance moves typical of social housing renters and the longer distance moves of the residual group who are mostly private sector renters).

To test ‘prototype’ HMA_s by comparing standardised house prices between the constituent areas requires the estimation of a hedonic regression model (Dale-Johnston, 1982). The details of this estimation are given in the Appendix but the price equation can broadly be written in algebraic terms as:

$$P = \alpha + \beta_1 S + \beta_2 T + \beta_3 D + \beta_4 R + \beta_5 M + \varepsilon$$

where

P	sale price of house;
S	structural attributes;
T	market conditions;
D	distance to major centre of population
R	residential density of neighbourhood
M	house type mix of neighbourhood

The choice of variables reflects both the nature of the task (in that it is not concerned with significant local neighbourhood effects) and the practicalities of a nationwide study. These regression models are inevitably subject to missing variable bias and this is reflected in our interpretation of the results.

5. Constructing a Tiered HMA Geography

The analysis considers different levels of closure for the different housing market tiers. A useful starting point in the choice of containment criteria is the 66.67% level used for defining TTWAs which are the official labour market areas in the UK. When this level of closure is applied to migration data then the areas produced are fewer – and so larger on average – than TTWAs (which were of course based on commuting data). This does not accord with the theoretical perspective that commuting-based *Framework HMAs* are either of a similar size to, or larger than, the migration-based *Local HMAs*. This point is also made by Hincks and Wong (2010). The way forward taken by this analysis is therefore for the definition of *Framework HMAs* to be based on higher levels of commuting closure than that used for TTWAs. This in fact puts into practice the theoretical principle that sees the *Framework HMAs* as wider labour market areas, defined so as to include longer distance commuters. As a result, lower levels of closure are deemed appropriate for the migration analyses to define *Local HMAs*.

Framework HMAs based solely on Commuting

The grouping analysis of commuting flows applies the same method and data that produced the 140 TTWAs in England but, with the higher 75% closure criterion needed to contain longer distance commuters, it produces 85 HMAs covering all of England (plus some adjacent parts of Scotland and Wales where there are strong flows across the border). Changing the closure criterion to 77.5% and applying it to all commuters produces 75 HMAs: this indicates that around this level there is only a modest level of sensitivity of the definitions to change in the key closure criterion. In general, this approach produces areas that perform well on technocratic criteria such as being fairly similar in size, and having few non-contiguities (nb. the grouping is not limited to contiguous area pairs because to do so prevents the process finding the optimal solution during its process of many thousand iterations). Using closure criteria that are around 75% produces large HMAs around metropolitan areas due to longer distance commuting, and this is appropriate for *Framework HMAs* because they reflect the impact and importance of high income households and their longer distance commuting on spatial arbitrage.

Local HMAs based solely on Migration

Migration closure defines a *Local HMA*. Applying the TTWA algorithm with the 66.67% closure criterion generates 86 HMAs. Reducing the migration closure criterion to 60% and 55% leads to increases in the number of HMAs to 152 and 223 respectively, whilst setting the closure level at 50% yields 327 HMAs. This indicates a steep level of sensitivity of the results to the closure criterion. There is also a very considerable difference in the size between the areas defined in the south and those in the old industrial regions. In some parts of the country there are also rather large numbers of non-contiguities. The use of migration data alone, with wards as the building block areas, thus tends to produce boundaries that do not perform ideally on the more technocratic criteria. That said, these problems are somewhat less at the 50% closure level so the *local HMAs* based on this criteria have more credibility, although at this level the northern conurbations do tend to be broken down into large numbers of areas.

Unitary HMAs based on combining Commuting and Migration

This approach combines analyses of commuting in a two stage grouping process that uses migration flows in its second step. It does not align with the theoretical tiered perspective but it provides useful insights into the relationship between commuting and migration. In this approach step 1 is identical to the commuting approach only, but is followed by a step 2 which takes account of migration patterns. The resultant geography requires the HMAs to be relatively self-contained in terms of both labour market areas (shown by commuting) and also migration. The basic method is not hierarchical. What this means is that step 1 labour market areas which do not pass the migration flow self-containment test are broken down into their constituent wards and these are then re-grouped on an individual basis so the final set of boundaries are as optimal as possible. Of course, the key decision remains the choice of just how self-contained the areas must be: to take one particular example, 72.5% commuting closure and 55% migration closure produced 93 HMAs. This is 8 more than the set defined with the slightly higher closure (75%) level applied to commuting the data alone (described above). Several of the 'additional' HMAs in the geography based on two datasets are in more peripheral areas, whilst there are some larger HMAs in more urban areas.

This approach does not fully accord with the theory of a tiered housing system. This is because the theory as elaborated earlier envisaged that migration-based areas would be either smaller than or a similar size to the commuting-based areas: by contrast, this approach has used migration data to define areas which are *larger* than those which were originally defined by reference to commuting flows. One justification for this approach could rest upon the fact that the theory was based on the study of urban economies, with restrictive assumptions then made to generalise the processes. When actual data on a complex mix of urban and rural areas are analysed it is not very surprising if some divergent patterns are found. More specifically, the commuting-based areas which are found here to not have 55% closure in terms of migration patterns are predominantly rural areas where there are more retired migrants (who are not subject to the constraints of commuting), and where there are numerous similarly sized settlements without a single dominant urban settlement around which all the flow patterns focus.

A Tiered Approach with Lower Level Areas based on Migration within Upper Level Areas based on Commuting

This approach seeks to follow the guidance of tiered HMA theory by defining commuting-based upper tier areas directly from individual wards and then subdividing these areas on the basis of migration self containment criteria. In this way both *Framework* and *Local HMAs* are established in one system. The algorithm first allows the upper tier boundaries to be more optimally defined based on the commuting criterion (step 1), but it creates a technical challenge because there is no existing method for disaggregating an upper tier set of areas into the largest possible number of lower tier areas that satisfy the migration self containment criteria (step 2). The technical innovation here is to treat then each of the *Framework HMAs* as a separate problem, so the second step of the analysis takes each area's constituent wards individually and then groups them until they meet the migration self-containment criterion *without* allowing any of these groupings to cross the *Framework HMA* boundaries which, as result, then form an upper tier boundary set.

Reflecting the evidence gathered from the earlier analyses summarised above, the closure criteria applied ranged from 75 to 80% and 50 to 60% commuting and migration respectively. There is a possibility that an upper tier area may not meet the criterion for lower tier migration closure, but with these criteria this did not happen. This does in fact occur with three of the smaller upper tier areas which are defined if the closure rates are set at the lower levels of the ranges (75% commuting and 50% migration): it is notable that these are all rural areas. A summary of the impact of changing the criteria is given in Table 1; ultimately the selection of the levels of closure is, as noted above, a purely empirical question.

Table 1 The effect on numbers of *Local HMAs* of nesting within *Framework HMAs*

% Commuting closure	Number of <i>Framework</i> <i>HMAs</i>	% Migration closure	Number of <i>Local HMAs</i> (if nested)	Number of <i>Local HMAs</i> (not nested)
75.0	85	50	280	327
75.0	85	55	204	223
75.0	85	60	135	152
77.5	75	50	280	327
77.5	75	55	200	223
77.5	75	60	130	152
80.0	58	50	285	327
80.0	55	55	204	223
80.0	53	60	131	152

Map 1 shows tiered HMAs defined by 77.5% commuting closure of *Framework HMAs* (step 1) and 50% migration closure for *Local HMAs* (step 2). The map shows where the major cities are through the location of the nine largest urban areas identified for research on the State of the English Cities (Champion 2006). Many of the more rural upper tier areas are not divisible at a lower tier, due to the factors such as the length of rural migration flows that have been noted above. As a result, a separate lower tier mainly applies to the more metropolitan parts of the country. For example the

Framework HMAs encompassing the provincial cities of Manchester and Liverpool both have 15 *Local HMAs* embedded within them.

These geographies are in nested tiers: the lower tier of *Local HMAs* is bounded by the limits of the upper tier *Framework HMAs*. Such a nesting approach can only reduce, not increase, the number which would be produced if the *Local HMA* geography is defined in an unconstrained way. The significance of this process is shown in Table 1. For example, constraining within the tiered structure reduces 50% *Local HMAs* from 327 to 280. This constraining effect may not be ideal from a technocratic viewpoint, but it is deemed here a relatively small price to pay to produce HMA geographies which accord with theoretical principles.

6. HMAs and Standardised House Prices

As explained earlier, the final stage in the process of developing appropriate HMA boundaries involved the experimental use of the boundaries within the hedonic price analyses described in the Appendix. The results of the analyses in general are that most of the prototype HMA geographies pass the pair-wise test that standardised house prices are statistically different. The largest number of similar contiguous pairs which fail this test is found when using the current TTWAs (Table 2).

Table 2 Comparison of Results of Hedonic Tests of prototype HMA Geographies¹

Geography	No. Spatial Units	No. Contiguous Pairs with same ² Standardised Price	% Reduction in Standard Error	Error Efficiency Index ³
Commuting 77.5% Closure	74	0	24.4	32.97
Commuting 75% Closure	82	1	25.3	30.85
TTWAs (66.7% Commuting Closure)	163	40	23.5	14.42
Nested Lower Tier 50% Migration Closure	277	7	29.5	10.66
Local Authorities	352	3	31.4	8.92

- 1 The number of spatial units do not sum to the totals in Table 1 because island areas have no contiguous areas and cannot be included in these tests.
- 2 The difference between the results for the two areas are not significantly different at the 5% level
- 3 % Reduction in standard error (x100) / number of spatial units in geography.

Map 1

Lower tier migration-based (50%) *within* commuting-based upper tier (77.5%)

The pairs of TTWAs concerned are shown in Map 2: it is notable that they are all in more peripheral and rural parts of England. This confirms there is a rurality issue, in the sense above, stemming partly from the differences in commuting and migration flows in these areas. TTWAs are smaller because their closure criterion was just 66.67% and in practice this results in a high proportion of TTWAs composed of just smaller towns with traditional market areas around them. Migration flows are often longer in rural areas, especially where there are more retired people, which results in spatial arbitrage operating across the boundaries of such TTWAs.

Map 2

Contiguous pairs of TTWA with similar standardised house prices (ie. not significantly different at the 5% level)

There are also more fundamental problems about the hedonic analysis and tests that arise partly because of the tiered nature of the housing market noted above and partly because of the difficulties in the use of the regression technique. It has been noted that there exist quite localised price differences between neighbourhoods (and wards) so the tests of pairs of proto-type HMAs may be dominated by these influences. The regression technique is also not powerful enough to account of this issue because it is not a fully specified model, partly because of missing variables. The model specification was driven by the availability of consistent variables on a national basis and is not designed to address potential neighbourhood premiums.

The hedonic analysis has been utilised to compare the efficiency of potential different HMA geographies. The localised nature of the housing market means that spatially disaggregated models should produce better results compared with a national model. The efficiency of the different geographies can be measured by a reduction in the standard error of the local regression models summed together in comparison of the national regression model. The results are given in Table 2. As expected the most localised geography considered, 352 local authorities (LAs) produces the greatest reduction in standard error, 31.4%. The most 'efficient' geography considered, defined by an index that looks at the percentage reduction in standard error per area in the geography, is the set of *Framework HMAs* (Map 1) derived from the analysis of commuting flows with 77.5% closure level (see column 5 in Table 3).

7. Conclusions and Implications

The theoretical perspective views the housing markets as a layered system characterised as follows.

Tier 1: Framework HMAs defined by long distance commuting flows and the long term spatial framework with which housing markets operate,

Tier 2: Local HMAs defined by migration patterns that determine the limits of short term spatial house price arbitrage

Tier 3: Submarkets defined in terms of neighbourhood or house type price premiums.

The argument then is that instead of a single tier geography of HMAs there should be different geographies reflecting the layers of the housing system. The analysis here has been concerned with developing HMAs that represent the top two tiers. The construction of the HMAs is based on the same grouping algorithm as that used to define TTWAs in the UK and several other countries.

Local HMAs based on migration patterns at the closure level of 50% have the greatest credibility on technocratic grounds as they have less non-contiguity problems but they do produce HMAs that vary considerably in size. HMAs in the northern conurbations in particular are very narrowly defined spatially, but this actually reflects very localised patterns of movement in these areas. The favoured geography for *Framework HMAs* is based on 77.5% commuting closure and produces HMAs much larger than TTWAs, the official local labour markets (whose closure level is 66.7%). The preference for the high level of closure is based on the role of the commuting patterns of high income earners in spatial arbitrage.

Spatial economic theory sees migration-based areas as either smaller than or a similar size to the commuting-based areas. It can be argued that the divergence from this pattern is confined to rural areas, and at least partly reflects the fact that theory was developed to model large cities and their hinterlands. For example, the areas produced by 75% closure of commuting areas which were found not to meet 50% closure in terms of migration were all predominantly rural areas where there are more retired migrants (who are not subject to the constraints of commuting), and where there are numerous similarly sized settlements without a single dominant urban settlement around which all the flow patterns focus. There were, in any case, no such problems with the final set of 77.5% closure *Framework HMAs*, because the

50% migration closure *Local HMAs* could all be defined within these upper tier boundaries. At these levels of migration/commuting closure then, the housing system is revealed in broad terms to be composed of a single tier in rural areas and two tiers in larger city regions. The upper tier of the *Framework HMAs* relates to a wide definition of commuting to take account of the housing market tentacles of long distance commuting by high income earners.

References

- Alonso W (1964) *Location and Land Use*, Harvard University Press, Cambridge.
- Briant A, Combes P-P and Lafourcade M (2010) Dots to boxes: do the size and shape of spatial units jeopardize economic geography estimations?, *Journal of Urban Economics*, 67, 287-302.
- Brown P and Hincks S (2008) A framework for housing market area delineation: principles and application, *Urban Studies*, 45, 2225-2247.
- Casado-Diaz JM Local labour market Areas in Spain: a case study, *Regional Studies*, 34, 843-856.
- Champion T (2006) *The changing urban scene: demographics and the big picture* Department for Communities and Local Government, London.
- CLG, Department for Communities and Local Government (2007) *Identifying sub-regional housing market areas: advice note*, Department for Communities and Local Government, London.
- Coombes M. (2009) English rural housing market policy: some inconvenient truths *Planning Practice and Research*, 24, 211-231.
- Coombes M (2010) Defining labour market areas by analysing commuting data: innovative methods in the 2007 Review of Travel-to-Work Areas, in J Stillwell, O Duke-Williams and A Dennett (Eds) *Technologies for migration and commuting analysis: spatial interaction data applications* Hershey, New York.
- Coombes M, Green A and Openshaw S (1986) An efficient algorithm to generate official statistical reporting areas: the case of the 1984 Travel-to-Work Areas revision in Britain *Journal of the Operational Research Society* 37, 943-953.
- Coombes M (2000) Defining locality boundaries with synthetic data, *Environment & Planning A*, 32, 1499-1518.
- Dahmann D (1999) New approaches to delineating metropolitan and nonmetropolitan settlement: geographers drawing the line, *Urban Geography*, 20, 683-694.
- Dale-Johnson, D (1982) An alternative approach to housing market segmentation using hedonic price data, *Journal of Urban Economics*, 11, 311-332
- Evans A W (1973) *The Economics of Residential Location*, Macmillan, London.
- Fox K and Kumar K (1965) The functional economic area: delineation and implications for economic analysis and policy, *Papers in Regional Science*, 15, 57-85.
- Green R, Malpezzi S and Mayo S K (2005) Metropolitan-specific estimates of price elasticity of supply of housing and their sources, *American Economic Review*, 95, 2, 334-339.

- Gyourko J (2009) Housing Supply, *Annual Review of Economics*, 1, 295-318.
- Hincks S and Wong C (2010) The spatial interaction of housing and labour markets: commuting flow analysis of North West England, *Urban Studies*, 47, 3, 620-649.
- Jones C (2002) The definition of housing market areas and strategic planning *Urban Studies*, 39, 549-654.
- Jones C, Leishman C and MacDonald C (2009) Sustainable urban form and residential development viability, *Environment and Planning A*, 41, 7, 1667 – 1690.
- Jones C and Watkins C (2009) *Housing Markets and Planning Policy*, Wiley-Blackwell, Oxford.
- Kan K (2002) Residential mobility with job location uncertainty, *Journal of Urban Economics*, 52, 3, 501-523.
- MacLennan D, Munro M and Wood G (1987) Housing choice and the structure of urban housing markets, in Turner B, Kemeny J and Lundquist L (Eds) *Between State and Market Housing in the Post-Industrial Era*, Almquist and Hicksell, Gothenburg.
- McMillen D and Smith S (2003) The number of sub-centres in large urban areas, *Journal of Urban Economics*, 53, 321-338.
- Muth R (1969) *Cities and Housing*. University of Chicago Press, Chicago.
- Openshaw S and Taylor P (1981) The modifiable areal unit problem, in N Wrigley and R Bennett (eds) *Quantitative Geography* Routledge, London.
- Pinto S M (2002) Residential choice, mobility, and the labor market, *Journal of Urban Economics*, 51, 3, 469-496
- Pryce G (1999) Construction elasticities and land availability: A two-stage least-squares model of housing supply using the variable elasticity approach, *Urban Studies*, 36, 13, 2283-2304.
- Schnare A and Struyk R (1976) Segmentation in urban housing markets, *Journal of Urban Economics*, 3, 146-166.
- Straszheim M. (1975) *An Econometric Analysis of the Urban Housing Market*, National Bureau of Economic Research, New York.
- Tolbert C and Killian M (1987) *Labor Market Areas for the United States*, Staff Report AGES870721, US Department of Agriculture.

Appendix

The testing procedure assessing differences in local house price structures is in three stages:

1. A hedonic price equation is parameterised for all transactions in the year in England;
2. Specific hedonic equations are estimated for each proto-type HMA.
3. Adjacent proto-type HMA's are paired and hedonic equations are estimated for the "pooled" subsample.

The purpose of this first stage is to control for property market heterogeneity and define a standardised house type for testing in stage two. As semi-log is a common form of such a model, we specify the dependent variable as the log of sale price. SPSS was used for the model estimation. For the third stage, the implicit price estimates in each HMA are compared with those of the "pooled" models. This final stage involves using a Chow Test which is employed to identify whether differential prices are observed for the standardised product in different markets. This procedure has been applied frequently to test for the existence of submarkets within HMAs but not before at the HMA level (see Jones and Watkins, 2009). It is unnecessary to test for 'parameter stability' for HMA's that are not physically attached.

The tests were applied to 960,000 transactions in 2005 for each different geography considered across England. The task required the construction of separate data sets for each HMA and each geography prior to the regression analysis. For each geography the comparative regression analysis typically involved estimating 700 equations.

Data and Variables

Land Registry transactions are the source of house price data. As the boundaries for the proposed test HMAs are known each house sale can be attributed to a HMA via its postcode. This data has been matched to the Census data to obtain a series of socio-economic including among other variables house type mix, household and dwelling density. Using GIS software additional location and neighbourhood variables have been added to Census Output Area (COA) level which are areas comprising an average of 125 households/300 people.

Table A1: Variables used in Hedonic Price Model

Variable Label	Description
LNPRICE	The natural logarithm of the house selling price
newhse	Dummy 1 for new-built house, 0 otherwise
d	Dummy 1 for detached houses, 0 otherwise
s	Dummy 1 for Semi-Detached houses, 0 otherwise
t	Dummy 1 for Terraced houses, 0 otherwise
nth2	Dummy 1 for sale in February, 0 otherwise
nth3	Dummy 1 for sale in March, 0 otherwise
nth4	Dummy 1 for sale in April, 0 otherwise
nth5	Dummy 1 for sale in May, 0 otherwise
nth6	Dummy 1 for sale in June, 0 otherwise
nth7	Dummy 1 for sale in July, 0 otherwise
nth8	Dummy 1 for sale in August, 0 otherwise
nth9	Dummy 1 for sale in September, 0 otherwise
nth10	Dummy 1 for sale in October, 0 otherwise
nth11	Dummy 1 for sale in November, 0 otherwise
nth12	Dummy 1 for sale in December, 0 otherwise
DetDom	Dummy 1 for COAs with > 50% detached houses, 0 otherwise
SemiDom	Dummy 1 for COAs with >50% semi-detached houses, 0 otherwise
TerDom	Dummy 1 for COAs with >50% terraced houses, 0 otherwise
FlatDom	Dummy 1 for COAs with >50% flats, 0 otherwise
Dens	Dwelling density - dwellings per hectare
Dist	Distance (km) to nearest town centre
Dist_Sq	Distance (km) to nearest town centre squared

The distance variable (Dist) is the distance to the centroid of the urban area within the HMA that had the highest population and is estimated using the

following procedure. Each of the sales in England in 2008 has a postcode reference. Data available through the academic web based data resource UK Borders allows access current and historical postcode indexes for the UK, this in turn gives the centroid coordinates for each UK postcode so we can accurately spatially locate the sale.

Within each HMA the dominant, in terms of population, urban area is identified. Using urban area polygons sourced from UK Borders applied 2001 Census data which urban area with the highest population is identified. Using the centroid of each of the highest populated urban areas in each HMA we then calculated the distance from each sale centroid to that urban area using Pythagoras.

In order to be able to examine the housing mix of each area, the available data on the percentage of each house type found in each of the COAs were converted. Dummy variables were created for COAs that are dominated (over 50%) by a specific house type. An additional dummy variable was created for a COAs that are not dominated by any house type. For density, the physical measure of dwelling per hectare (in a specific COA) was used, fitting better the data than other (highly correlated to this variable) population density.