## Housing Supply, Housing Demand, and Affordability

Bernard Fingleton

*Urban Stud* 2008; 45; 1545
DOI: 10.1177/0042098008091490

Urban Studies
45(8) 1545–1563, July 2008

# Housing Supply, Housing Demand, and Affordability

**Bernard Fingleton**

*[Paper first received, June 2006; in final form, August 2007]*


## Abstract

The affordability of housing is a major policy issue that has increasingly become a concern for UK government as house prices have risen dramatically in recent years. This is partly because of the importance of affordability for the recruitment and retention of key workers, many of whom are on national pay scales and earning salaries that do not fully reflect the differences in prices that exist, in particular between London and the South East and the rest of Great Britain. Government policy is to increase the supply of housing in order to improve affordability in the greater South East. However, assuming that this expansion in housing supply is also to be accompanied by an expansion in employment, the outcome is that there will be both an increase in supply and in demand for housing, with the counter-intuitive result that, under one of the scenarios set out in this paper, in some areas affordability will worsen rather than improve.


## Introduction

This paper is about the consequences of changing the supply of housing within the greater South East of England, as advocated in recent UK government policy. Housing supply and its implications for affordability in England have recently been considered in an ambitious modelling exercise by Meen *et al.* (2005) and this provides a background to the present paper. The topic has also recently been given some impetus from a theoretical standpoint by the work of Glaeser *et al.* (2005, p. 2), who observe that "the modern literature on urban growth and economic geography generally ignores housing supply". It is from this starting-point that they develop an attractive conceptual framework by which to examine the consequences of a shock to the housing supply in the North American urban context. From a UK housing policy perspective, the report by Wilcox (2003) is also highly relevant, since it highlights the issue of the lack of affordability, particularly among so-called key workers in the public sector, who are increasingly finding that salaries are

**Bernard Fingleton** is in the Department of Economics, the University of Strathclyde, Sir William Duncan Building, 130 Rottenrow, Glasgow, G4 0GE, United Kingdom. Fax: 0141 548 4445. E-mail: bernard.fingleton@strath.ac.uk.

0042-0980 Print/1360-063X Online

DOI: 10.1177/0042098008091490

falling below what is necessary for owner-occupation. His detailed spatial analysis is parallelled in the approach of this paper, given the significant variation in affordability across small areas. Recent developments in UK government policy with regard to housing reflect this problem, with proposals for a sharp increase in the supply of housing in particular parts of the greater South East region of England in the period up to 2015. This paper explores possible implications of this 'exogenous shock' to the UK system, but it does not closely follow the line of analysis pursued either by Wilcox (2003), Glaeser *et al.* (2005) or by Meen *et al.* (2005). Wilcox (2003) presents some interesting methodology, but also ignores the issue of spatial interaction which is fundamental to the simulations in this paper. Likewise Glaeser *et al.* (2005), in the interest of simplicity, abstract from the effects of commuting and ignore spatial interaction. In the context of the current work, explicit consideration of both of these facets is an indispensable element of the analysis.

The complex multi-equation modelling system of Meen *et al.* (2005) does incorporate spatial interaction, although in many other regards the present paper differs from their approach. For instance, the basic spatial unit in Meen *et al.* (2005) is the Government Office Region, of which there are 9 covering England and Wales, compared with the 353 local authority districts used in the present paper. Also, their estimating equations involve time and hence necessarily involve numerous other variables, such as the mortgage interest rate and rate of growth of the FTSE index, and, although in both approaches house prices are determined by the interactions of supply and demand, the similarity ends there. Additionally, their suite of models includes intricate demographic sub-models, with interactions between demographics and housing and labour markets. In contrast, the purely cross-sectional modelling approach adopted here is much simpler in construction.

Also, the present paper uses contemporary theory base on Dixit–Stiglitz theory of monopolistic competition, so that internal increasing returns in the producer services sector of the urban economy drive increasing returns to employment density, leading to higher wage levels in dense central cities. This theory-driven approach is very dissimilar to the more eclectic labour market modelling in Meen *et al.* (2005) in which, for instance, average earnings depend on various region-specific time-series including feedback from house prices. Glaeser *et al.* (2005) also approach wage determination somewhat differently. They assume that residents experience a common level of utility across all areas, which is the balance of wage, house price and amenity and local public good differences between areas. Hence the assumption is that lower wages and/or higher house prices are compensated by a higher level of amenity/public goods with no loss of utility, likewise either higher wages and/or lower prices compensate for lower amenity.

With this background, in the paper I use simulation methods to examine the impact of the UK government's policy with respect to the supply of housing in the South East of England (OPDM, 2005). The provision of more homes might in general be expected to reduce house prices. However, the simulations I report in the paper support an alternative thesis, that an indirect consequence of increased housing supply will be less rather than more affordable housing, at least in some parts of the greater South East. A similar phenomenon occurs with road building; it is often the case that more roads induce more traffic and worsen congestion. In addition, there will evidently be some negative environmental consequences. Both these conclusions have major policy implications and so have not been stated without qualification or arrived at without careful consideration of the assumptions and methods employed. These are provided in detail in what follows.

Part 2 of the paper introduces a spatial econometric house prices model in order to explain year 2001 price variations across 353 small areas in England—namely, Unitary Authority and Local Authority Districts, or UALADs. In part 3, some preliminary initial simulations of house prices are reported, which rest on the assumption that the estimated model coefficients remain constant but the level of housing supply changes in a way that is broadly consistent with avowed UK national and local government policy. Apart from some hopefully intelligent guesswork as to the precise location of the additional homes, this simulation effort is preliminary because it is carried out without any consideration of commensurate changes in demand and yet we know that, under the UK government's proposals, extra housing will change housing densities and employment is expected to grow 'alongside' these extra homes. The consequences for house prices are considered in part 4.

The analysis of parts 2, 3 and 4 ignores changes to wage levels, which will affect both the level of demand and affordability, which is defined as an area's mean house price divided by the mean annual wage level available from employment in the area. To accommodate wage variations, part 5 introduces the second spatial econometric model, in this case to explain wage levels across areas in the year 2000. As noted earlier, a central feature of the wages model is the assumption that there are increasing returns to employment density, reflecting the greater efficiency of production. The extra employment associated with the expansion of housing will therefore, under this model, change wage rates and this will have an impact on demand, which is determined jointly by employment levels and by wage levels. However, an increase in wage levels will also to some extent offset the rise in prices arising from the stimulus to demand and therefore affordability will be improved. Part 6 describes the resulting changes to house prices and affordability; part 7 concludes the paper, re-emphasising the conditional nature of the analysis.

## The House Price Model

### Theory

In order to determine the house price level that follows from an exogenous positive increment to the supply of houses as a result of government policy, we need to take account of changes to the level of demand ($q_j$). The assumption here is that housing demand responds to changing wage levels and employment levels, both locally and within commuting distance. Housing demand from within the local area is simply a function of income from local jobs, equal to the local wage rate ($w$) times the local employment level ($E$).

Housing demand due to wage and employment levels within commuting distance of $j$ is assumed to be equal to

$$w_j^c E_j^c = \sum_k \exp(-\delta_j D_{jk}) w_k E_k, \quad j \neq k, D_{jk} \leq 100km \tag{1}$$

Equation (1) indicates that jobs located within 100 km of area $j$ contribute to total income with a weight determined by the area-specific exponent $\delta_j$ and by the distance between residential areas $j$ and $k$ ($D_{jk}$), with $\delta_j$ being estimated using observed census data on travel-to-work patterns.[1] Table 1 shows the overall proportion of workers travelling various distances to work from home, with $\delta_j$ ($j$ = 1 ... 353) obtained from 353 individual commuting tables similar to Table 1. This means that area $j$'s exponent $\delta_j$ is determined by its commuting data, as a

**Table 1.** Commuting distances in Great Britain (2001 census)

| | *Commuting distance in km* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | *<2* | *2–5* | *5–10* | *10–20* | *20–30* | *30–40* | *40–60* | *≥60 km* |
| Percentage | 23.21 | 23.34 | 21.19 | 17.66 | 6.20 | 2.73 | 2.52 | 3.14 |

result of iterating the function exp $(-\delta_j D_{jk})$ through a range of values to obtain the value giving the closest fit, defined as the minimum of the sum of the squared deviations of the observed proportions in each distance band up to 60 km and the proportions of the sum of the function exp $(-\delta_j D_{jk})$ calculated using the upper limit of each distance band. Thus, the individual estimation of $\delta_j$ allows for differential propensity to commute by location.

Housing demand $(q_j)$ is also negatively related to the price of housing $(p_j)$. Given that high prices drive down demand, it is assumed that high prices 'nearby' will cause demand otherwise attributable to nearby locations to be displaced, spilling over into $j$. We refer to this as a displaced demand effect. Hence it is assumed that demand at $j$ will be positively related to the weighted average of prices in surrounding areas, which is denoted by the matrix product $\mathbf{W}p_j$. The $n$ by $n$ matrix $\mathbf{W}$ is a row standardised version of matrix $\mathbf{W}^*$, Hence for cell $(j, k)$

$$W_{jk} = \frac{W^*_{jk}}{\sum_k W^*_{jk}} \qquad (2)$$

with $W^*_{jk} = \dfrac{1}{d^2_{jk}}$, in which $d_{jk}$ is the straight-line distance between locations $j$ and $k$, and $W^*_{jk} = 0$ for $d_{jk} \geq 50$ km.

It seems reasonable to assume that the spillover does not extend very far, since often market knowledge is localised and market conditions change significantly with distance, so we approximate the localised interaction by assuming that it only involves areas less than 50 km apart and falls quite sharply as distance increases.

Demand also depends on variables such as amenity and local public goods, referred to collectively as $A_j$, and on other unmodelled factors such as demand coming from non-wage-earners such as the retired and students, and the effects of criminality, social quality of the neighbourhood, local taxes, etc., which are represented by a stochastic error $\omega \sim N(0, \sigma^2 I)$. Hence the demand function is

$$\begin{aligned} q_j &= a_0 + a_1 w_j E_j + a_2 w^c_j E^c_j \\ &\quad + a_3 A_j - a_4 p_j + \nu \mathbf{W} p_j + \omega \end{aligned} \qquad (3)$$

The supply function is

$$q_j = b_0 + b_1 p_j + b_2 O_j - \eta \mathbf{W} p_j + \varsigma \qquad (4)$$

which assumes, *ceteris paribus*, that the level of housing supply $q_j$ increases in the price at $j$. For example, property-owners in high-price areas may be more likely to want to realise the value of their assets by offering to sell; in contrast, in low-price locations homeowners may prefer to withhold their properties from the market. Likewise, property developers will be attracted to areas with high prices and, by the same token, it is also assumed that high prices nearby $(Wp_j)$ will attract supply away from $j$, hence the negative sign for $\eta$. This is referred to as a displaced supply effect.

In addition, controlling for price effects, supply also is assumed to relate to the size of the existing stock of properties $(O)$, which is represented by the number of owner-occupier households reported in the 1991 Census of Population.[2] This usefully pre-dates the current period and can be considered to be predetermined with respect to 'current' prices. The stock will affect the quantity of housing supplied to the market. Large cities will naturally supply more properties to their market than do small villages, given that in a large city there will be many more property-owners; in other words, $O$ will be larger. In the simulations that follow, it is assumed that the supply of housing is altered due to some areas' stock $O$ being changed as a result of government policy.

The reduced form is obtained by normalising the supply function with respect to $p$, thus

$$p_j = \frac{1}{b_1} q_j - \frac{b_0}{b_1} - \frac{b_2}{b_1} O_j - \frac{\eta}{b_1} \mathbf{W} p_j + \frac{\varsigma}{b_1}$$

and substituting for $q$ gives

$$p_j = c_1[a_0 + a_1 w_j E_j + a_2 w_j^c E_j^c + a_3 A_j - a_4 p_j + v \mathbf{W} p_j + \omega] - c_0 - c_2 O_j - c_3 \mathbf{W} p_j + \xi$$

Simplifying and introducing the number of square km per household $A_S$, the square of the distance of the area from London $A_L$, and the level of educational attainment $A_E$ (see Appendix) in place of $A_j$ gives

$$p_j = \rho \sum_{k \neq j} W_{jk} p_k + d_0 + d_1 w_j E_j + d_2 w_j^c E_j^c + d_3 A_{Ej} + d_4 A_{Sj} + d_5 A_{Lj} + d_6 O_j + \varepsilon_j \tag{5}$$

Writing equation (5) in general matrix terms, with columns of the $n$ by $k$ matrix $\mathbf{X}$ comprising the right-hand-side variables apart from the spatial lag, gives

$$p = (I - \rho \mathbf{W})^{-1} (\mathbf{X} d + \varepsilon) \tag{6}$$

in which the errors $\varepsilon_j \sim N(0, \tau^2 I)$. allow for measurement error in the price variable and for other unmodelled effects.

The displacement effects in the supply and demand functions combine to give the term $\rho \mathbf{W} p_j$ in the reduced form (5), assuming that the matrix $\mathbf{W}$ is common to both. The restriction of this assumption is somewhat lessened by virtue of the separate parameters $v$ and $\eta$, which combine as the parameter $\rho$ in the reduced form. This indicates a direct spatial interaction between prices in $j$ and in surrounding areas. One way in which the existence of these spillover effects, which are crucial to the outcome of the simulations, can be tested, is to set $\rho$ equal to zero and examine the consequences using standard diagnostics for residual spatial autocorrelation. The diagnostics do indicate very significant spatial autocorrelation among the regression residuals (see Table 2).

## Estimates of the Price Model

The price data $p_j$ are the overall average selling prices (all property types) by UALAD for the year 2001, as provided by the Land Registry. Using these data, Table 2 gives the results of fitting the house price model (5) via OLS, ML and 2sls. In the OLS estimation, all the right-hand-side variables are assumed to be exogenous and there is no spatial lag. In the ML estimation, the spatial lag is present and by definition endogenous, but the other right-hand-side variables are exogenous. In the 2sls estimation, both the spatial lag and $A_E$ are assumed to be endogenous.

Table 2 shows that the residuals for the OLS estimated model are highly spatially autocorrelated, pointing to a misspecified model, due perhaps to one or more of omitted spatially autocorrelated variables, simultaneous equation bias or a nuisance spatial error process. It turns out that what appears to be important is the omission of the endogenous spatial lag as proposed under equation (5).

The ML estimates in columns 4 and 5 show that the endogenous spatial lag is highly significant, indicating that prices in area $j$ affect, and are affected by, prices 'nearby'.

**Table 2.** Estimates of house price models (dependent variable $p$)

| | *OLS* | | *ML* | | *2sls* | |
|---|---|---|---|---|---|---|
| | *Parameter estimate* | *T–ratio* | *Parameter estimate* | *T–ratio* | *Parameter estimate* | *T–ratio* |
| Constant | –505505.98 | –6.08 | –587509.75 | –8.83 | –663715.83 | –5.85 |
| $wE$ | 786.50 | 9.68 | 385.29 | 5.72 | 298.89 | 3.91 |
| $w^cE^c$ | 46.37 | 10.96 | 16.58 | 4.06 | 9.68 | 2.08 |
| $O$ | –0.5329 | –5.24 | –0.2338 | –2.83 | –0.1635 | –1.86 |
| $A_E$ | 160644.78 | 7.48 | 161626.69 | 9.42 | 177105.79 | 6.08 |
| $A_S$ | 374005.79 | 3.15 | 295210.48 | 3.11 | 263968.53 | 2.68 |
| $A_L$ | –0.3674 | –6.77 | –0.1067 | –2.26 | –0.0502 | –0.99 |
| $\rho$ | — | — | 0.6300 | 12.99 | 0.7624 | 12.29 |
| $R^2/\bar{R}^2$ | 0.6260 | | 0.7614 | | 0.7642 | |
| Standard error | 39141.35 | | 31252.73 | | 31401.28 | |
| Log likelihood | –4230.3372 | | –4165.39 | | — | |
| Residual correlation | I = 13.49 | | LM = 0.7251 | | Z = –0.6496 | |
| Degrees of freedom | 346 | | 345 | | 345 | |

*Notes*: $\bar{R}^2$ = squared correlation actual and fitted. For the OLS model, we use the conventional $R^2$ statistic. LM is distributed as $\chi^2_1$ under the null hypothesis of no residual spatial autocorrelation. I is the standardised value of Moran's I-statistic for residual spatial autocorrelation. Z is the standardised value from the Anselin–Kelejian (1997) statistic for residual spatial autocorrelation with a spatial lag. The spatial autocorrelation tests use the matrix **W** defined in the text.

Also, house prices increase significantly with increasing local demand ($wE$) and with increasing demand within commuting distance ($w^cE^c$) and that they fall significantly with increasing stock ($O$). Also, prices increase significantly as a result of better schooling locally ($A_E$), when there is more space per household ($A_S$), and when distance from London ($A_L$) is less.

Stated in more detail, a unit increase in 'schooling' ($A_E$) causes estimated expected house prices to rise by £161 627. Since the key stage 2 values range from 3.547 to 4.280, this indicates that moving from the lowest achieving area to the highest increases house prices by about £118 473. Similarly, more physical space per household ($A_S$) is estimated to cost about £295 210 per square kilometre. Since the variable $A_L$ is the square of km distance, then the negative impact on price of each extra km is greater the further the location is away from the capital, being equal to –0.1067 times twice the distance.[3]

On the supply side, the indication is that 1000 extra houses cause prices to fall by £234 (which is additional to the fall in price caused by the loss of amenity caused by the greater housing density).

The 2sls estimates are based on first-stage regressions with regressands defined as the endogenous spatial lag and the 'endogenous' $A_E$, and exogenous regressors $wE$, $w^cE^c$, $O$, $A_S$ and $A_L$, their exogenous first spatial lags obtained via the matrix products of the row-standardised matrix $\mathbf{W}$ and $wE, w^cE^c, O, A_S$, and $A_L$, giving $\mathbf{W}wE$, $\mathbf{W}w^cE^c$, $\mathbf{W}O$, $\mathbf{W}A_S$ and $\mathbf{W}A_L$, and 46 county dummy variables (coded 1 if the UALAD was within a county, zero otherwise, and eliminating Tyne and Wear to avoid the dummy variable trap). The resulting estimates are given in columns 6 and 7 of Table 2. However, it turns out that treating $A_E$ either as endogenous or as exogenous makes very little difference to the results, and a standard test of exogeneity due to Hausman, shows that exogeneity is a reasonable assumption for the purposes of statistical estimation. Given that $\mathbf{W}p$ is endogenous, we test whether $A_E$ is endogenous by looking at a function of the change in the coefficient on $A_E$ estimated by two-stage least squares under either assumption for $A_E$ (see Maddala, 2001). Using the same instruments as for Table 2, the relevant test statistic is equal to 0.05643 which has a very large tail probability (pr. = 0.8122) in the $\chi^2_1$ reference distribution, suggesting that an exogeneity assumption for $A_E$ is appropriate.

### Simulations Based on Increasing Housing Supply

In the simulations, I show the effect on prices and affordability of increasing the level of the housing stock in selected UALADs. Affordability was defined earlier in terms of mean prices and wages, but there are alternative measures, as set out by Meen *et al.* (2005). Mean *et al.* (2005, p. 15) focus on the ratio of lower quartile house prices to incomes,[4] but note that "the main econometric modelling is conducted in terms of mean mix-adjusted house prices. Long time-series of quality-adjusted median and lower quartile prices are not available". In their set-up, to avoid such data limitations, the assumption is that the mean, median and lower quartile house prices retain a constant relationship.

The choice of UALADs is based on the government's own stated policy to increase the supply of housing in the south and east of England. In the ODPM's *Sustainable Communities: homes for all* (OPDM, 2005), there are four growth areas, the Thames Gateway and East London, Milton Keynes/South Midlands, London/Stansted/Cambridge/Peterborough and Ashford in Kent. The understanding from this documentation is that the Thames Gateway will provide approximately 120 000 homes by 2016 and the other three about 300 000 new homes by 2016.

These number are of course to some extent informed guesswork and they are subject to approval and confirmation by the Regional Spatial Strategy process, but they provide an approximate indication of the magnitude of the expansion that is anticipated. In order to carry out the simulation at a UALAD-specific level, it is necessary to allocate these new homes to UALADs using information from local planning authorities. The uncertainties involved and lack of easily accessible information make this somewhat difficult to do with any real precision across all UALADs, although as an illustration I have attempted a more detailed allocation for the Cambridge sub-region that is based on currently existing planning documents and policy statements, although, as with the other allocations of additional homes across the other areas, the outcome is a very provisional set of numbers.

In the case of the region around Cambridge, I have used the Memorandum by Cambridgeshire County Council (SHC 26, printed 14 January 2003) put before the Select Committee of the Office of the Deputy Prime Minister: Housing, Planning, Local government and the Regions, which indicated the future housing needs for the Cambridge sub-region for the period 1999–2016. This gives the following allocation of new housing

Within the built-up area of Cambridge 9 000
On the edge of Cambridge 8 000
New settlement at Oakington/Longstanton 6 000
In market towns and main rural centres 17 000
Elsewhere in sub-region 8 000

These 48 000 new homes were allocated to specific UALADs by studying the strategic options for the Cambridge sub-region examined in the report commissioned for Cambridgeshire County Council, Go-East and EEDA (*Implementing the Cambridge sub-regional strategy*). The outcome of this analysis is that the housing growth needs were allocated as follows

Cambridge 9 000
South Cambridgeshire (edge of city) 8 000
Longstanton 6 000
1889 in each of 9 market towns/rural centres 17 000
Others[5] 8 000

With regard to the 8000 homes split between the areas, these were initially allocated equally across seven UALADS (excluding Cambridge) within the Cambridge sub-region. However, the growth area designated by the ODPM (ODPM, 2005, Figure 3b) excludes Forest Heath and St Edmundsbury so their homes were reallocated to East Cambridgeshire and South Cambridgeshire respectively.

In terms of UALAD allocations, this finally gives

Cambridge 9 000
East Cambridgeshire 6 064
Fenland 3 032
Huntingdonshire 6 810
South Cambridgeshire 20 064
Uttlesford 3 032

While these allocations are to some extent a matter of conjecture, they are considered to be within the constraints to be found within the various planning guidelines and documents consulted and in line with the expansion of housing proposed by ODPM.

For the other areas of expansion, the method used is very much cruder. For the 13 areas of the Thames Gateway, omitting any expansion in the City of London, the remaining 12 areas have had the total of 120 000 homes equally divided, giving 10 000 new houses for each area. The local authority areas each receiving 10 000 additional homes are therefore Barking and Dagenham, Bexley, Dartford, Greenwich, Hackney, Tower Hamlets, Havering, Lewisham, Newham, Redbridge, Thurrock and Walton

Forest. For the remaining areas—that is, those outside the Thames Gateway and not in the Cambridge sub-region—there are 300 000–48 000 = 252 000 houses to be allocated. As a first approximation in this process of allocation, this number has simply been divided equally between the 34 areas, giving 7412 new homes in each. Figure 1 shows the allocation of new homes across UALADs. Adding these additional homes to the original variable $O_j$ gives the new values $O_j^S$ which are used in the simulations.

Let us assume that the number of jobs and their location, and their wage rates, stay the same as in 2000. This is not a realistic assumption, but is simply a way of isolating the effect of the increase in housing supply *per se*. The new levels of housing stock ($O_j^S$) are therefore combined with the original values for local demand ($wE$), demand within commuting distance ($w^cE^c$), schooling ($A_E$), space per household ($A_S$) and distance from London ($A_L$). I also use the residuals from the fitted model ($\hat{\varepsilon}$) as an additional variable. The assumption is that the omitted variables captured by these residuals will remain constant. Therefore, the equation used to obtain the simulated values is

$$p^S = (I - \hat{\rho}\mathrm{W})^{-1}(\mathrm{X}^S\hat{d} + \hat{\varepsilon}) \qquad (8)$$

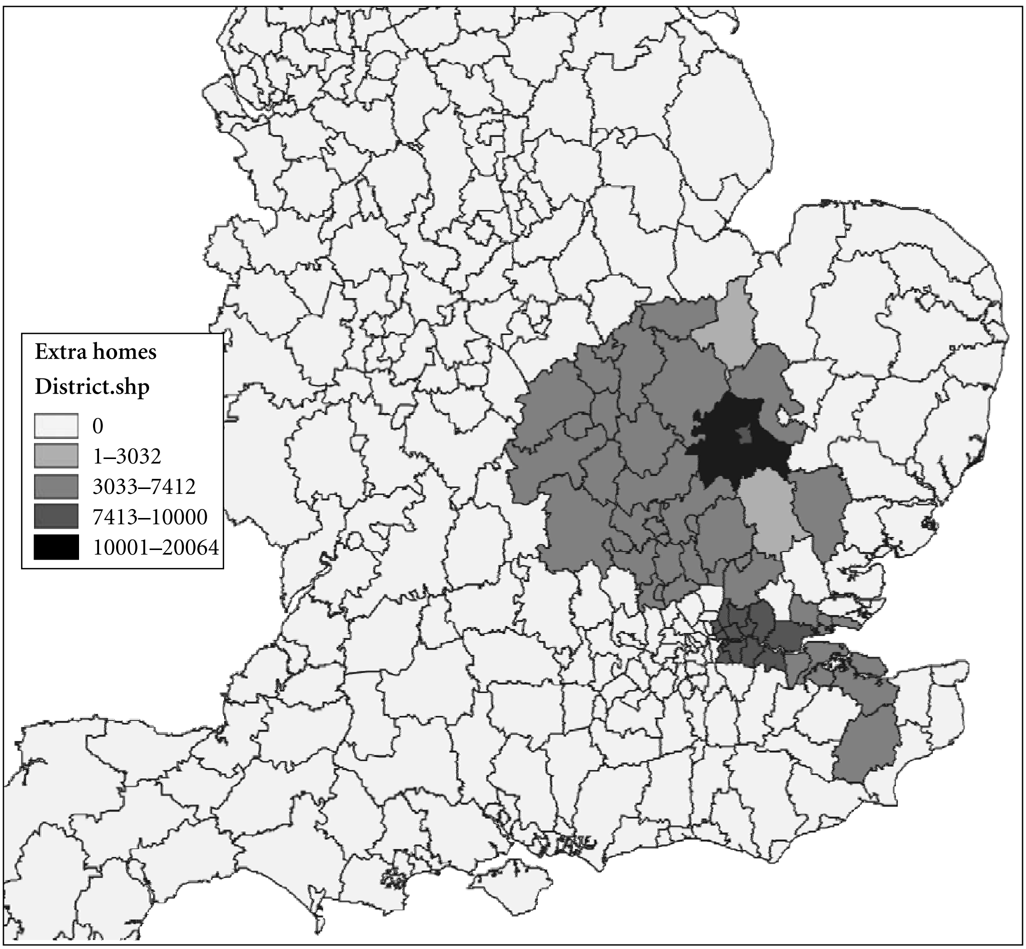


**Figure 1.** The conjectured allocation of homes

where, $X^S$ is the $n$ by $k$ matrix of regressors similar to $\mathbf{X}$ in equation (6) but with $O^s$ replacing $O$; $\hat{\rho}$ and $\hat{d}$ are the maximum likelihood coefficient estimates given in column 4 of Table 2, and $\hat{\varepsilon}$ are the residual estimates produced by these coefficients.

The outcome is that house prices fall and affordability rises by the same proportion. This is apparent from Figure 2, which shows the proportionate change in house prices (and affordability) across the region. In Cambridge city, the impact is a 3.44 per cent fall in house prices, compared with 2 per cent in Hackney in Inner London.

Figure 3 shows the additional negative impact on prices of increased residential density. This is achieved by replacing the original variable space per household $A_S$ by $A_S^S$, which is equal to the square kilometre area of each UALAD divided by the simulated stock of homes $O^S$, again obtaining simulated prices using the Table 2 ML coefficient estimates. The outcome is a greater reduction in prices and a bigger increase in affordability. For example, in South Cambridgeshire, prices fall by 7.08 per cent as a result of the combined effect of the increased supply of homes and the negative impact of higher

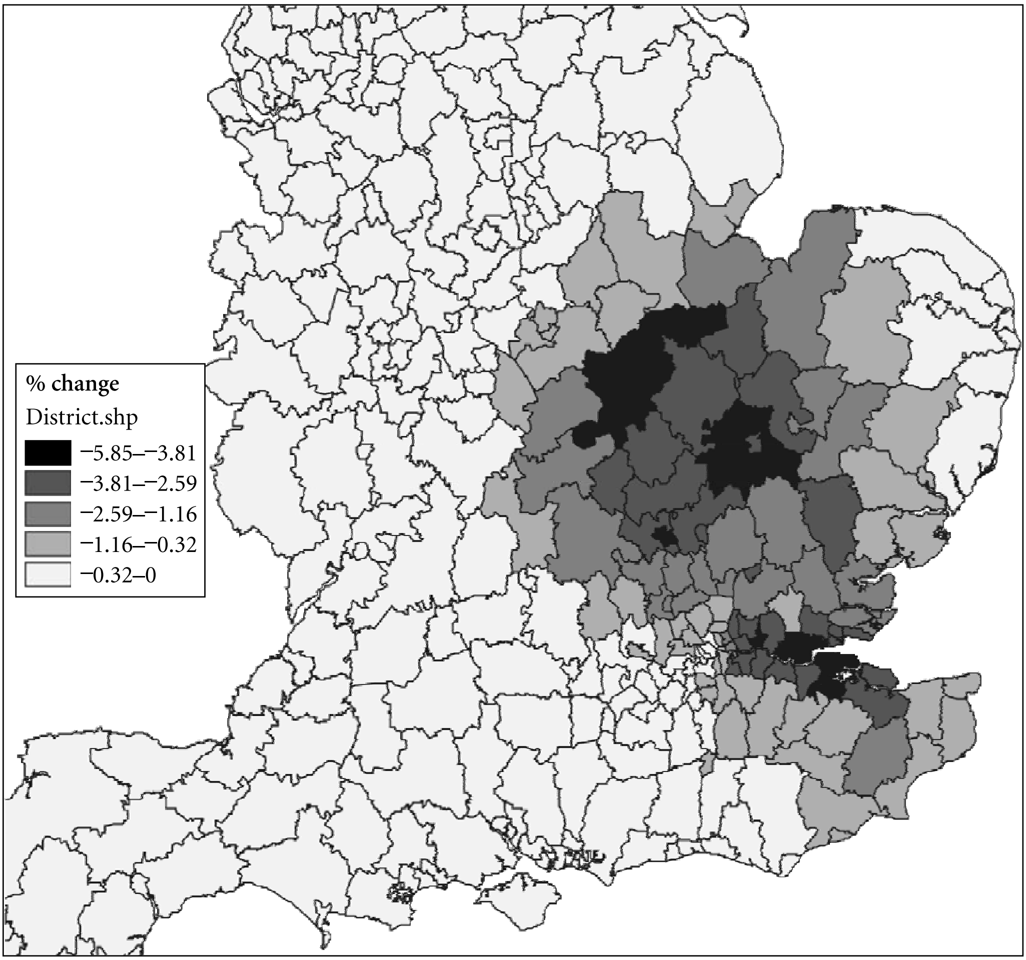


**Figure 2.** The percentage change in prices and affordability (initial estimate) as a result of the increased number of homes

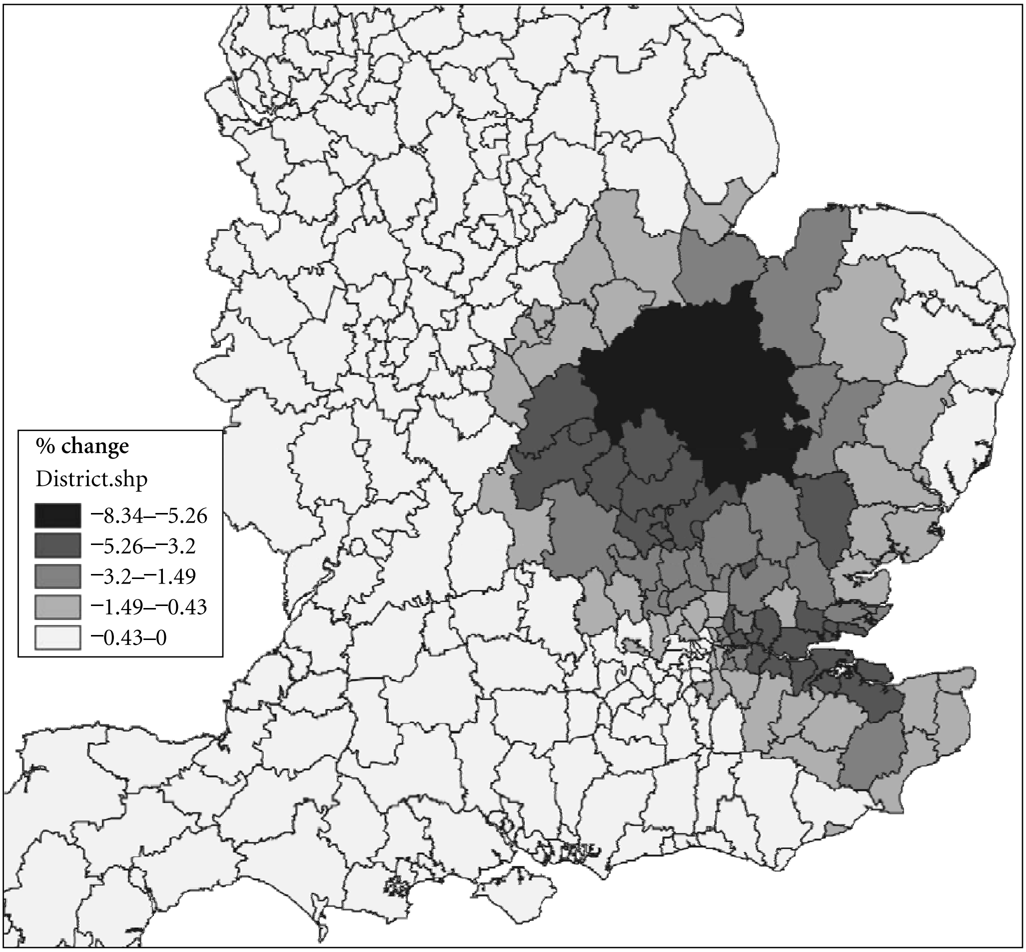


**Figure 3.** The percentage change in prices and affordability as a result of the increased number of homes and amenity loss due to extra density

density on amenity, compared with 4.95 per cent simply as a result of the increased housing supply. In Cambridge city and Hackney, the price changes inclusive of loss of amenity are –4.54 per cent and –2.07 per cent.

## Introducing Employment Changes

The assumption thus far has been that new homes are constructed but there is no change in the number or location of jobs. However, it appears more reasonable to assume that the number of jobs and possibly their locations will change. With regard to the number of jobs, an initial scenario is that total employment grows in each area at its long-run growth rate, obtained for the period 1971–2003 for each UALAD, and is thus entirely independent of what is assumed for housing. Summing across UALADs within the Greater South East, the total number of net new jobs (since employment falls in some areas) is equal to 756 372.

The second scenario is under the assumption that there will be some spatial reallocation of this total, which is distributed among UALADs to be within commuting distance of their new residential development described earlier. In order to allocate these extra jobs, we make a preliminary allocation of a proportion of the net new jobs to each UALAD according to its proportion of all additional homes ($O_j^S - O_j$). Then the allotted new jobs $\tilde{J}_j$ are redistributed on the basis of the commuting propensity specific to each UALAD, using the formula

$$\bar{J}_{jk} = \tilde{J}_j \frac{e^{-\delta_j D_{jk}}}{\sum_k e^{-\delta_j D_{jk}}} \qquad (9)$$

in which the $\delta_j$ are exactly the same ones as used for equation (1). The final number of new jobs for UALAD $j$ is

$$\bar{J}_j = \sum_k \bar{J}_{jk}$$

For example, of the preliminary assignation of $\tilde{J}_j = 13348$ jobs to Luton, some of these are allocated via equation (9) to other areas within commuting distance of the new Luton homes. Also, some jobs assigned to other areas at the preliminary stage are redistributed to Luton via equation (9), although the net outcome is a final total of $\bar{J}_j = 8212$ and $E_j^S = E_j + \bar{J}_j$. In reality, a process of dual causation is more likely than jobs simply following homes or vice versa. In practice, the simulation could equally well have commenced with an allocation of employment to areas such that housing located within commuting distance, perhaps according to the pattern of Figure 1. Despite the way this has been done, no direction of causation is assumed.

## The Wages Model

### Theory

The analysis up to this point has assumed that wage levels will remain the same as the number of homes expands, but it is more reasonable to assume also that wages may change as housing supply expands and new employment is created. The question is, do homes become more affordable once we allow for changing wage rates? For example, if we assume that, as a consequence of the additional homes and jobs, wages rise, then this itself fuels demand, while at the same time higher wage rates will make houses more affordable.

The assumption of a relationship between wage rates and employment derives from the theory described by Rivera-Batiz (1988), Abdel-Rahman and Fujita (1990), Ciccone and Hall (1996) and Fujita and Thisse (2002), which relies on pecuniary externalities as a result of producer services under a monopolistic competition market structure, internal increasing returns and assuming a constant elasticity of substitution production function for producer services. These services are an input, in the form of the level of composite services ($I$), into the production function of competitive industry, which values variety of service inputs. The competitive industry production function is

$$Q = (E_c^{\beta} I^{1-\beta})^{\alpha} L^{1-\alpha} \qquad (10)$$

in which, $\alpha$ and $\beta$ are parameters; $E_C$ denotes employment in competitive industry, and $L$ is the area of land in which production takes place.

We assume production per unit area, so that $L = 1$ and therefore $Q' = (E_c^{\beta} I^{1-\beta})^{\alpha}$. Given that the only input assumed for producer services is labour, the reduced form from this theory is a relationship between competitive industry output ($Q'$) and employment density ($E'$ is total employment per unit area) so that

$$\ln Q' = \ln \phi + \gamma \ln E' \qquad (11)$$

in which, $\phi$ and $\gamma$ are two parameters to be estimated. With increasing returns, one would expect $\gamma$ to exceed 1 in value (see Fujita and Thisse 2002, p. 102).

Assuming an equilibrium allocation of production factors, the coefficient $\alpha$ is the share of $Q'$ that goes to $E'$ (rather than the other factor $L$) and, using standard equilibrium theory in which the wage rate equals the marginal product of labour, one obtains

$$w = \frac{\alpha Q'}{E'} \tag{12}$$

and substituting this into equation (11) gives

$$\ln w = \ln(\alpha\phi) + (\gamma - 1)\ln E' \tag{13}$$

This means that, provided $\gamma$ is greater than 1, wage rates increase with the density of employment.

### The Wages Model Estimates

The wages model given in equation (13) is very simple and does not allow for the many other factors affecting compensation levels. One important factor is the variation in worker efficiency across areas and, to allow for this, we work in terms of labour efficiency units, so that the wage equation becomes

$$\ln w = \ln(\alpha\phi) + (\gamma - 1)\ln(SE') \tag{14}$$

in which $S$ denotes the level of efficiency.

Fingleton (2003, 2006) argues that, if we assume that efficiency levels spill over between areas due to commuting, then a possible model for efficiency levels is

$$\begin{gathered} \ln S = \mathbf{X}\mathbf{b} + \delta\tilde{\mathbf{W}}\ln S + \tilde{\varepsilon} \\ \tilde{\varepsilon} \sim N(0, \Omega^2) \end{gathered} \tag{15}$$

in which, $\mathbf{X}$ is an $n$ by $k$ matrix of exogenous variables; $\mathbf{b}$ is a $k$ by 1 vector of coefficients; the matrix product $\tilde{\mathbf{W}}\ln S$ is an $n$ by 1 vector with coefficient $\delta$; and $\tilde{\varepsilon}$ represents unmodelled variables that behave as random shocks.

The endogenous variable $\tilde{\mathbf{W}}\ln S$ is the contribution to efficiency of workers working in area $j$ that is derived from commuting, as defined by the matrix $\tilde{\mathbf{W}}$. In this case, we assume that

$$\tilde{W}_{jk} = \frac{e^{-\kappa_j D_{jk}}}{\sum_k e^{-\kappa_j D_{jk}}}$$

In the estimation of the parameters $\kappa_j$, precisely the same method is used as described for equation (1). However, the set of commuting matrices is commuting from work rather than home, as used by Fingleton (2003). At the aggregate level, the matrix is practically identical to Table 1 and therefore is omitted to save space; however, at the local level there are differences. For example, workers in the City of London have a much shallower distance decay than do residents.

Fingleton (2003, 2006) shows that substituting equation (15) into equation (14) and simplifying leads to the following reduced form

$$\begin{aligned} \ln w = {} & \delta\tilde{\mathbf{W}}\ln w + k_1 + (\gamma - 1) \\ & (\ln E' - \delta\tilde{\mathbf{W}}\ln E') + \mathbf{X}\mathbf{c} + \Psi \\ & \Psi \sim N(0, (\gamma - 1)^2\Omega^2) \end{aligned} \tag{16}$$

in which, $k_1$ is a constant since $\tilde{\mathbf{W}}$ is row-standardised and $\mathbf{c} = (\gamma - 1)\mathbf{b}$.

This is then estimated by a complicated iterative routine that satisfies the constraints on the coefficients. In this paper, the method is greatly simplified by making the assumption that the weighted average of labour efficiency levels in 'surrounding' areas $\tilde{\mathbf{W}}\ln S$ can be approximated by the weighted average of surrounding area wage rates $\tilde{\mathbf{W}}\ln w$; hence, $\ln S = \mathbf{X}\mathbf{b} + \delta\tilde{\mathbf{W}}\ln w + \tilde{\varepsilon}$, in which case the reduced form is

$$\begin{aligned} \ln w = {} & (\gamma - 1)\delta\tilde{\mathbf{W}}\ln w + k_1 \\ & + (\gamma - 1)\ln E' + \mathbf{X}\mathbf{c} + \Psi \end{aligned} \tag{17}$$

The matrix **X** comprises a column of 1s, and two exogenous determinants of efficiency variations. One is the natural logarithm of the relative concentration in 1999 by UALAD of employees in the computing and research and development sectors (*lnT*). The relative concentration is given by the location quotient based on employee totals by UALAD for 1992 SIC 72 and 1992 SIC 73 taken from the annual business enquiry employee analysis (available through NOMIS). These SIC codes relate to employees in hardware consultancy, software consultancy and supply, data processing, database activities, computer and office machinery maintenance and repair, other unspecified computer-related activities, natural sciences, engineering, social sciences and humanities research. The second covariate (*lnF*) is the log of the percentage of residents with no qualifications given by the UK's 2001 Census.[6] The rationale for this variable is the widely recognised link between labour inefficiency and inadequate schooling. Although $F$ post-dates $w$ by one year, it is assumed to be exogenous for the purposes of estimation. It is assumed that it is unlikely that there would be feedback from $w$ to $F$ on this time-scale, and $F$ is also likely to be affected by factors other than wage differentials, such as institutions and social and cultural differences, and by government and EU policy initiatives, and so it seems appropriate to treat it as predetermined. Comparing the 2001 and 1991 shares with no qualifications for the 408 UALADs of Great Britain, while the average population share with no qualifications has fallen dramatically, a strong linear correlation ($r = 0.872$) exists between the 1991 and 2001 census datasets, so using the 1991 data gives similar results.

Written out in full and simplifying the parameterisation, the wages model that is estimated is

$$\begin{aligned}\ln w = k_1 + \lambda\tilde{\mathbf{W}}\ln w + (\gamma - 1)\ln E' \\ +\tau \ln T + \pi \ln F + \Psi\end{aligned} \tag{18}$$

which is the well-known spatial lag model.

The observed wages $w$ are taken from the year 2000 results of the Office for National Statistics' New Earnings Survey, which is carried out annually by the UK's Office of National Statistics. These are workplace-based survey data of gross weekly pay for male and female full-time workers irrespective of occupation. These are available on the NOMIS website (the Office for National Statistics' on-line labour market statistics database). There are no data for the Isles of Scilly, so the data for the nearest mainland area of Penwith have been used in this case.

The ML estimates for the wages model are given in Table 3. The estimated elasticity on employment density $\gamma - 1$ indicates that doubling city density causes wages to rise by about $\ln(2^{0.04}) = 2.8$ per cent, which is somewhat lower than the estimates typically obtained elsewhere in the literature, which take no account of spatial interaction. Ciccone (2002) estimates an elasticity of approximately 0.045 for productivity with respect to the density of economic activity using data on European regions and, according to the literature survey by Rosenthal and Strange (2004, p. 2133), "doubling city size seems to increase productivity by an amount that ranges from roughly 3–8 per cent". The coefficient estimates for the labour efficiency variables $\ln T$ and $\ln F$ are significant and appropriately signed, and the spatial lag coefficient $\lambda$ is also significantly different from zero. To highlight the significance of the spatial effect, also given are the OLS estimates of the model

$$\begin{aligned}\ln w = \ln(\alpha\phi) + (\gamma - 1)\ln E' \\ +\kappa \ln T + \pi \ln F + \zeta \\ \zeta \sim N(0, \sigma^2)\end{aligned} \tag{19}$$

The diagnostics point to a significant spatial lag (LM lag greatly exceeds LM error) and it appears that, simply by making wage rates depend on wage rates in nearby UALADs,

**Table 3.** Estimates of wages models (dependent variable ln$w$)

| | *OLS* | | *ML spatial lag* | |
|---|---|---|---|---|
| | *Parameter estimate* | *T–ratio* | *Parameter estimate* | *T–ratio* |
| Constant | 6.242<br>(6.110) | 54.04<br>(51.59) | 4.256768 | 9.91 |
| ln$E$ | 0.04741<br>(0.04324) | 11.55<br>(10.37) | 0.040105 | 9.64 |
| ln$T$ | 0.06853<br>(0.07828) | 7.18<br>(8.07) | 0.049419 | 5.05 |
| ln$F$ | –0.2054<br>(–0.1579) | –5.80<br>(–4.28) | –0.161836 | –4.58 |
| city | (0.468) | (3.84) | | |
| $\kappa$, $\lambda$ | | | 0.323 | 4.823 |
| $\bar{R}^2$ | 0.580<br>(0.596) | | 0.6137 | |
| Standard error | 0.114<br>(0.112) | | 0.109407 | |
| Log–likelihood | 266.7528<br>(274.0634) | | 278.539192 | |
| Residual correlation (LM) | | | 0.003177 | |
| LM (error) | 37.69 | (29.97) | | |
| LM (lag) | 69.87 | (77.37) | | |
| Z | 6.544 | (5.852) | | |
| Z[1] | 6.860 | (6.139) | | |
| Degrees of freedom | 349 (348) | | 348 | |

*Notes*: $\bar{R}^2$ = squared correlation actual and fitted. Z is the standardised value of Moran's I-statistic for residual autocorrelation. Z[1] is the standardised value using the moment estimates from the empirical randomisation distribution (with 100 replications) of Moran's I for regression residuals. LM is distributed as $\chi^2_1$ under the null hypothesis of no residual spatial autocorrelation. The spatial autocorrelation test uses the standardised matrix $\tilde{\mathbf{W}}$ defined in the text. Figures in paretheses are the equivalent OLS estimates resulting from the omission of the most influential case, the City of London, from the dataset. The measure of influence is Cook's statistic, which for the City of London is equal to 14.6412, due to a high positive residual and a high leverage.

the residual pattern is eradicated, as evident from the small value for the LM statistic for the spatial lag model.

## Simulating Affordability: The Two Scenarios

Figure 4 gives affordability growth under the first scenario—that is, assuming no spatial reallocation of employment. The simulated price $p^S$ is based on variables $w^S$, $O^S$, $A^S_S$ and $E^S$ using equation (8), in which $w^S$ is the simulated wage given by

$$\ln w^S = (I - \lambda \tilde{\mathrm{W}})^{-1}(\mathrm{X}^S \hat{f} + \hat{\Psi}) \qquad (20)$$

which is the matrix equivalent of equation (18) but with ln$E'^S$ (the logarithm of projected employment density) in place of ln$E'$. Hence, the matrix $\mathbf{X}^S$ is the $n$ by $k$ matrix with columns equal to 1, $\ln E^{\prime S}$, $\ln T$ and $\ln F$, which means that we are assuming that labour efficiency remains constant over the period. The estimated coefficient $\hat{\lambda}$ and vector of coefficient estimates $\hat{f}$ are as in Table 3 and $\hat{\Psi}$ are the residual estimates produced by these coefficients.

The simulations taking account of wage rate variations mean that the positive growth in house prices does not equate to the negative growth of affordability. Actual affordability is 2001 price per UALAD divided by annual wage level earned from jobs located in each UALAD in 2000. Therefore, percentage affordability change is 100 $[\ln(p^S/w^S)–\ln(p/w)]$, with a positive value indicating lower affordability. In Cambridge, wages increase by 0.93 per cent and house prices fall by 1.04 per cent, so that affordability changes by –1.97 per cent. In South Cambridgeshire, the respective values are 1.43 per cent, –2.61 per cent and –4.04 per cent; while in Hackney, falling wages (–0.05 per cent) are overtaken by faster falling house prices (–4.09 per cent) so that affordability also improves by 4.04 per cent. Overall, the pattern is as one might anticipate, with improved affordability in areas where we have assumed the supply of housing will increase.

Figure 5 is the outcome under the second scenario—that is, assuming that the spatial distribution of employment change is

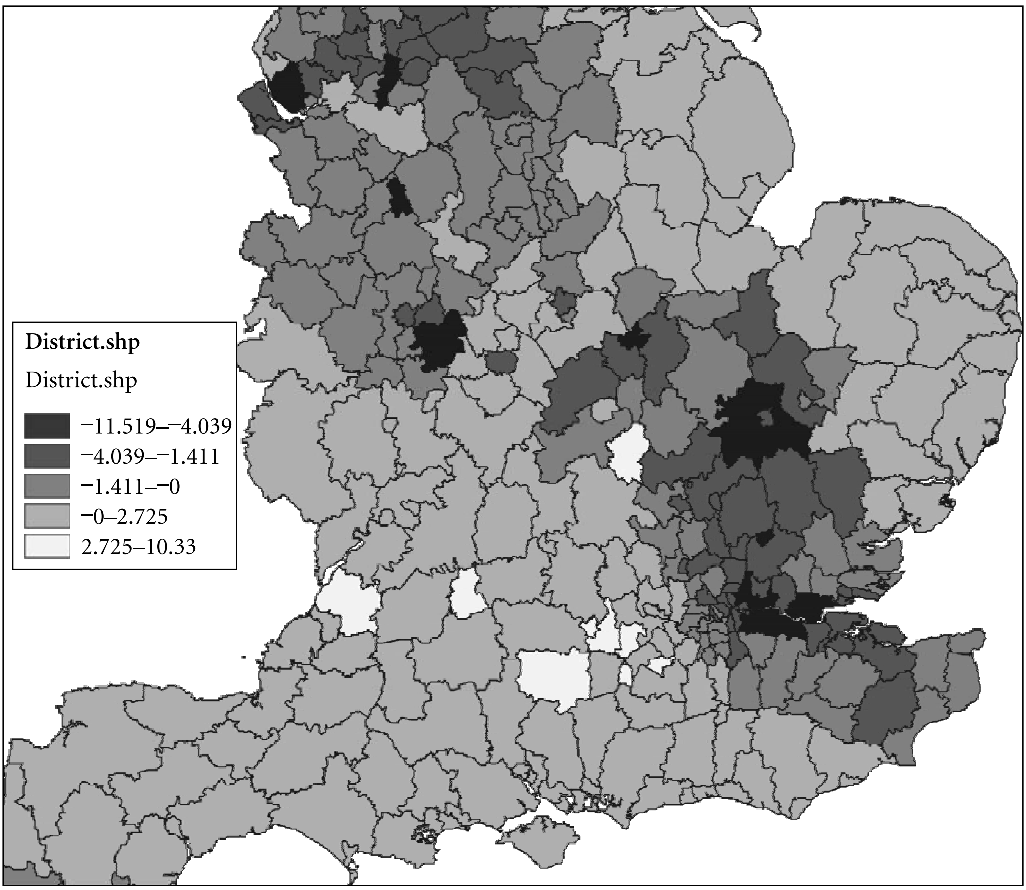


**Figure 4.** The percentage change in affordability using projected employment change with no spatial reallocation

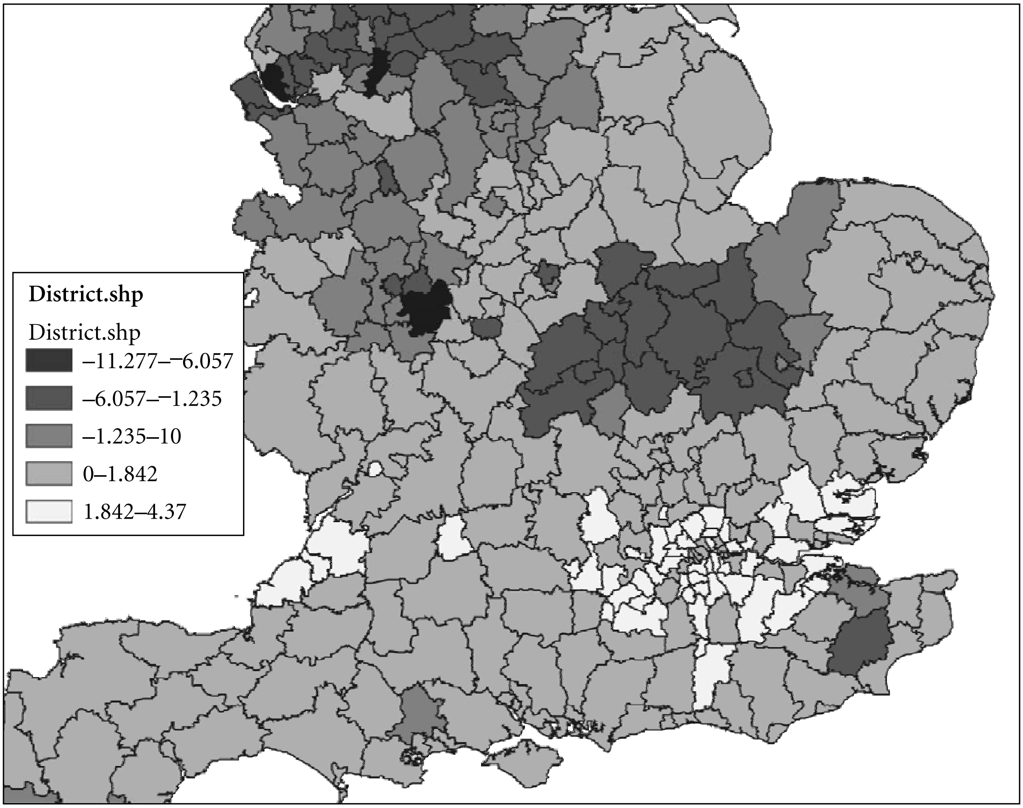


**Figure 5.** The percentage change in affordability using spatially reallocated projected employment change

moderated according to the conjectured location of new homes and commuting distances to these new homes. This now creates rising rather than falling employment in the Thames Gateway, which is what one might reasonably expect under the comprehensive regeneration that is planned for this area. The outcome in the Thames Gateway is now mainly worsening affordability. For example, in Hackney, the average price (all types of property) to (annualised) wage ratio goes from 6.3666 to 6.4783, commensurate with a loss of affordability of 1.74 per cent since, although rising employment density produces wages growth equal to 0.52 per cent, this is insufficient to offset 2.26 per cent house price growth. Despite the loss of affordability, in general the Thames Gateway remains a wedge of relatively affordable housing by comparison with many parts of inner London and the outer suburbs. In contrast to the Thames Gateway, in Cambridge affordability improves, changing by –2.19 per cent, since wages grow by 0.91 per cent and prices by –1.28 per cent; and in South Cambridgeshire affordability improves more, changing by –4.20 per cent as the net outcome of 1.02 per cent growth in wages and –3.18 per cent fall in prices. However, here the price to wage ratio is much higher than in Hackney, equal to 8.4298 in 2001 and 8.0828 in 2015.

## Conclusions

The analysis shows that there is a basis for assuming that the ODPM's plans to expand the supply of housing in the greater South East could in many areas produce less affordable rather than more affordable homes.[7] It also shows that a spatially disaggregated modelling perspective provides useful insights that are not available when working at a more aggregate level. For example, the spatial econometric modelling highlights the presence of localised knock-on effects, with house prices and affordability changing within an area as a result of policy intervention in other areas. Affordability is taken to be the ratio of average price to wage level in each UALAD; however, it is acknowledged that this is a very crude concept, since in reality there will be a distribution of wages and of prices within each area. Also, if house prices fall, and allowing for wage changes housing becomes more 'affordable', this may be partly due to an increase in supply, but part of the fall in price will (in most of the simulations) be due also to an increase in density and hence a fall in quality. It is inescapable that what is meant in arguing that housing may in places be more affordable is that a more affordable product is offered, with a lower price partly reflecting a lowering of quality due to a higher density.

When affordability falls, which as we see from our simulations is anticipated for much of the Thames Gateway, workers working within the area may prefer to commute in from lower-price areas. Likewise, residents will often seek jobs outside where wage rates are higher. Therefore, a likely consequence of less affordability is more in- and out-commuting. In contrast, an increase in affordability is likely to have more environmentally benign consequences, as the impetus to commute will be less.

Overall, therefore, we see from this paper that—on the basis of assumptions about the causes of house price variation and about the causes of wage variation, and on the basis of an assumed allocation of new homes and jobs to areas—a case can be made for the counter-intuitive outcome that more homes means, for many areas, less-affordable homes. While these assumptions are open to challenge, the purpose of the paper is to open up this debate, simply offering a first-round justification of the assumptions and models that have been employed. What have been presented here are not forecasts. The limitations of the data, the uncertainties associated with the assumptions, are sufficient to dispel that notion. There is not space here for a detailed test of the robustness of the conclusions to different assumptions. The thesis of this paper, which we re-emphasise should be the subject of further detailed analysis, is that the policy proposed by the ODPM, of expanding housing in the greater South East, is not axiomatically guaranteed to produce more affordable homes everywhere, and could possibly have some environmentally damaging consequences, with an increase in commuting as well as a fall in amenity levels, as measured by the negative impact of density on house prices, across a wide area of the South East of England.

## Notes

1. 2001 Census of Population, using NOMIS Table S120. The data are based on all people aged 16–74 who were working in the week before the Census. The distance travelled is a calculation of the straight line between the postcode of place of residence and postcode of workplace.
2. Local Base Statistics, Table L20 Tenure and amenities: Households with residents; residents in households. This is available in the NOMIS database.
3. Based on $\delta P/\delta V = -\beta = -0.1067$, and $V = d^2$, hence $\delta P/\delta d = -\beta 2d$.
4. This ratio was of particular interest to UK government as defining targets to be met by increasing the housing supply.
5. Others split between the areas (excluding Cambridge city) within the Cambridge sub-region according to judgement.
6. Available from the website Casweb, which is a web interface to statistics and related information from the UK Census of Population.
7. This somewhat counter-intuitive effect is not unanticipated by Meen *et al.* (2005), who note that the effect of an increase in new construction on affordability could be mitigated by increased

housing demand, although it is given much less emphasis than in the present paper.

## References

Abdel-Rahman, H. and Fujita, M. (1990) Product variety, Marshallian externalities and city size, *Journal of Regional Science*, 30, pp. 165–183.

Anselin, L. (1988) *Spatial Econometrics, Methods and Models*. Dordrecht: Kluwer.

Anselin, L. and Kelejian, H. H. (1997) Testing for spatial error autocorrelation in the presence of endogenous regressors, *International Regional Science Review*, 20, pp. 153–182.

Cheshire, P. and Sheppard, S. (2004) Capitalising the value of free schools: the impact of supply constraints and uncertainty, *Economic Journal*, 114(499), pp. F397–F424.

Ciccone, A. (2002) Agglomeration effects in Europe, *European Economic Review*, 46, pp. 213–227.

Ciccone, A. and Hall, R. E. (1996) Productivity and the density of economic activity, *American Economic Review*, 86, pp. 54–70.

Fingleton, B. (2003) Increasing returns: evidence from local wage rates in Great Britain, *Oxford Economic Papers*, 55, pp. 716–739.

Fingleton, B. (2006) The new economic geography versus urban economics: an evaluation using local wage rates in Great Britain, *Oxford Economic Papers*, 58(3), pp. 501–530.

Fujita, M. and Thisse, J.-F. (2002) *Economics of Agglomeration*. Cambridge: Cambridge University Press.

Gibbons, S. and Machin, S. (2003) Valuing English primary schools, *Journal of Urban Economics*, 53, pp. 197–219.

Glaeser, E. L., Gyourko, J. and Saks, R. E. (2005) *Urban growth and housing supply*. Discussion Paper No. 2062, Harvard Institute of Economic Research.

Haining, R. J. (1990) *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge: Cambridge University Press.

Leech, D. and Campos, E. (2003) Is comprehensive education really free? A case-study of the effects of secondary school admissions policies on house prices in one local area, *Journal of the Royal Statistical Society A*, 166, pp. 135–154.

Lucas, R. E. (1976) Econometric policy evaluation: a critique, in: K. L. Brunner (Ed.) *The Phillips Curve and Labor Markets*, pp. 19–46. Supplement to the *Journal of Monetary Economics*.

Maddala, G. S. (2001) *An Introduction to Econometrics*. New York: Wiley.

Meen, G., Allmendinger, P., Andrew, M. *et al.* (2005) *Affordability targets: implications for housing supply*. ODPM, London.

ODPM (Office of the Deputy Prime Minister) (2001) *Regional Planning Guidance for the South East.* RPG 9. London: HMSO.

ODPM (2005) *Sustainable communities: homes for all.* The Office of the Deputy Prime Minister, London.

Rivera-Batiz, F. (1988) Increasing returns, monopolistic competition, and agglomeration economies in consumption and production, *Regional Science and Urban Economics*, 18, pp. 125–253.

Rosenthal, S. and Strange, W. C. (2004) Evidence on the nature and sources of agglomeration economies, in: J. V. Henderson and J.-F. Thisse (Eds) *Handbook of Regional and Urban Economics Vol. 4*, pp. 2119–2172. Amsterdam: North-Holland.

Wilcox, S. (2003) *Can work – can't buy: local measures of the ability of working households to become home owners*. Report prepared for the Joseph Rowntree Foundation, York.

## Appendix

Several studies show that higher residential property prices correlate to a good local school (for example, Leech and Campos, 2003; Cheshire and Sheppard, 2004; Gibbons and Machin, 2003). There is a range of data on school performance at both primary and secondary level, but in this paper I choose pupils in the final year of primary school since they are in the crucial year for determining their secondary school destination; hence I use the results of the 1998 key stage 2 tests taken by 11-year-old pupils. The individual results have been converted by Oxford University into an indicator of the quality of schooling available to residents of very small administrative areas (wards). In this paper, to allow compatibility across variables, I have calculated the mean of the mean scores per ward, averaging over wards within each UALAD. The construction commences with 8413 (English) wards each with a (mean) score, and from these I have calculated the mean score for the 353 English UALADs (there are on average 24 wards per UALAD), to give the variable $A_E$ used in the regressions.