STRATHCLYDE

DISCUSSION PAPERS IN ECONOMICS

SIMULATING WAGES AND HOUSE PRICES USING THE NEG

BY

BERNARD FINGLETON

NO. 09-13

DEPARTMENT OF ECONOMICS
UNIVERSITY OF STRATHCLYDE
GLASGOW

Simulating Wages and House Prices Using the NEG*

Bernard Fingleton**

April 2009

Abstract
The paper incorporates house prices within an NEG framework leading to the spatial distributions of wages, prices and income. The model assumes that all expenditure goes to firms under a monopolistic competition market structure, that labour efficiency units are appropriate, and that spatial equilibrium exists. The house price model coefficients are estimated outside the NEG model, allowing an econometric analysis of the significance of relevant covariates. The paper illustrates the methodology by estimating wages, income and prices for small administrative areas in Great Britain, and uses the model to simulate the effects of an exogenous employment shock.

Introduction

The reduced form of the basic NEG model described by Fujita, Krugman and Venables[3] (1999) comprises a small number of simultaneous equations, typically with each region's economy characterized by two sectors, one under monopolistic competition and the other perfectly competitive. Typically, particularly at the international level, the sectors are industry and agriculture. However a major element of expenditure and consumption is housing, and this goes largely unrecognised in the standard representation of the model. However there have been several attempts to link housing to some version of the NEG model, notably by Helpman(1998), Hanson(2005) and Brakman, Garretsen and Schramm (2004). In their approaches the price of housing services is treated as a purely endogenous outcome related to income within the NEG model. In contrast, this paper introduces some other covariates, in addition to income, to create an ancillary model of house prices. One advantage of this approach is that variables in the housing submodel could be used to show the impact of exogenous factors. An additional feature of our modelling approach is that, within the core NEG model, and unlike many other applications, the location of economic activity is represented by labour efficiency units, following FKV(Ch. 15). In addition, the model assumes the presence of a spatial equilibrium (Hanson, 2005, Glaeser, 2008), in which the house price to disposable income ratio is constant across localities.

Basic Theory

Typically in NEG theory we have two different sectors, one under monopolistic competition (the *M* sector) and one under perfect competition (the *C*) sector, hence utility (*U*) depends on *M* and *C* thus

$$U = M^{\alpha} C^{1-\alpha}$$
$$0 \leq \alpha \leq 1$$

(1)

---

[3] Hereafter FKV.

in which α is equal to the expenditure share of M goods. The quantity M is given by the constant elasticity of substitution (CES) subutility function with an elasticity of substitution between any pair of varieties is σ, hence

$$M = \left[\sum_{i=1}^{x} m(i)^{(\sigma-1)/\sigma}\right]^{\sigma/(\sigma-1)} = \left[\sum_{i=1}^{x} m(i)^{\frac{1}{\mu}}\right]^{\mu} \qquad (2)$$

in which $m(i)$ denotes variety i, there are x varieties.

Following the normalizations given in FKV, five simultaneous non-linear equations comprise the reduced form of the empirical model, equations (3) and (4) for M and C wages (w_i^M and w_r^C), equations (5) and (6) for M and C prices (G_i^M and G_i^C), and equation (7) for wage income (Y_{1r}). Additionally, as shown by equation (8), nominal M wages and the M and C price indices determine real M wages (ω_i). In order to give quantitative values to these equations, we need values for the elasticities of substitution σ and η for M and C varieties respectively, we need to know λ_r and ϕ_r which are the respective shares of the total supply of M and C workers for $r = 1\dots R$, and we have to assign a value to the coefficient of the Cobb-Douglas preference function α.

$$w_i^M = [\sum_r Y_r (G_r^M)^{\sigma-1} T_{Mir}^{\,1-\sigma}]^{\frac{1}{\sigma}} \qquad (3)$$

$$w_i^C = [\sum_r Y_r (G_r^C)^{\eta-1} T_{Cir}^{\,1-\eta}]^{\frac{1}{\eta}} \qquad (4)$$

$$G_i^C = [\sum_r \phi_r (w_r^C T_{Cir})^{1-\eta}]^{\frac{1}{1-\eta}} \qquad (5)$$

$$G_i^M = [\sum_r \lambda_r (w_r^M T_{Mir})^{1-\sigma}]^{\frac{1}{1-\sigma}} \qquad (6)$$

3

$$Y_i = \alpha \lambda_i w_i^M + (1-\alpha) \phi_i w_i^C \tag{7}$$

$$\omega_i = w_i^M (G_i^M)^{-\alpha} (G_i^C)^{\alpha-1} \tag{8}$$

In the most simple case C goods and services are assumed to incur no transport costs, so that $T_{Cir} = 1$ and $w_i^C = 1$ across all i. These assumptions might be considered unrealistic but we can easily relax them (see FKV, chapter 7), as evident from equations (1…7). We can also allow the C sector to exhibit diversity while still remaining competitive (along the lines of the assumptions related to the Armington elasticity used in GCE modelling, which is equivalent to the elasticity of substitution η). This means that to operationalize the model by solving equations (3…7), there is a need to define M and C sectors, and then obtain values for the exogenous terms T_{Cir} and T_{Mir}, λ_r and ϕ_r, σ and η and the expenditure share α.

A simplified version with no C sector expenditure

Assume that $\alpha = 1$, so the C sector carries no utility and accounts for no expenditure share. This certainly seems a reasonable assumption in an urban setting where C denotes agriculture. In other words, C does not exist. This means that utility is given by

$$U = M^\alpha C^{1-\alpha} = M = \left[\sum_{i=1}^{x} m(i)^{\sigma-1/\sigma} \right]^{\sigma/\sigma-1} = x^{\sigma/\sigma-1} m(i) \tag{9}$$

$$\sigma > 1, x \to \infty$$

This simplifying assumption means that the problem of defining the two sectors M and C is avoided. In particularly it avoids the problem of identifying which goods and services posses no internal increasing returns to scale. Instead we assume that all firms have both fixed costs and variable costs and incur transport costs. We therefore start from the proposition that the M sector, involving returns to scale and transport costs, describes

all goods and services in the urban economy. Under this assumption, we only require one elasticity of substitution, σ, and one trade cost function T_{Mir}.

With $\alpha = 1$ the simultaneous equations become

$$w_i^M = [\sum_r Y_r (G_r^M)^{\sigma-1} T_{Mir}^{1-\sigma}]^{\frac{1}{\sigma}} \qquad (10)$$

$$G_i^M = [\sum_r \lambda_r (w_r^M T_{Mir})^{1-\sigma}]^{\frac{1}{1-\sigma}} \qquad (11)$$

$$Y_i = \lambda_i w_i^M \qquad (12)$$

$$\omega_i = w_i^M (G_i^M)^{-1} \qquad (13)$$

Introducing labour efficiency units

The equilibrium wage, price index and income levels calculated within the NEG equations (10…13) take no account of other factors affecting wages levels. We assume that the major omission is the level of efficiency of workers in different locations (see FKV, p. 264). Hence rather than labour units λ_r, we work with labour efficiency units equal to $\kappa_r \lambda_r$ in which κ_r is the level of efficiency of labour in region r. Therefore $\kappa_r \lambda_r$ is the number of labour efficiency units. Accordingly the simultaneous equations become

$$w_i^M = [\sum_r Y_r (G_r^M)^{\sigma-1} T_{Mir}^{1-\sigma}]^{\frac{1}{\sigma}} \qquad (14)$$

$$G_i^M = [\sum_r \kappa_r \lambda_r (w_r^M T_{Mir})^{1-\sigma}]^{\frac{1}{1-\sigma}} \qquad (15)$$

$$Y_i = \kappa_i \lambda_i w_i^M \qquad (16)$$

$$\omega_i = w_i^M (G_i^M)^{-1} \qquad (17)$$

In these w_r^M is the wage per efficiency unit of labour, ω_i is the real wage rate in efficiency units. It follows that $\kappa_r w_r^M$ is the wage per unit of labour and $\kappa_i \omega_i$ is the real wage rate per unit of labour.

Disposable income rather than wages

Owning an asset such as a house adds to credit worthiness in the form of collateral to be set against borrowing. We wish to take account of the fact that disposable income includes both borrowing and other income sources[4] and how this affects endogenous outcomes. The model outcomes we simulate would occur given the existence of a spatial equilibrium whereby the house price to disposable income ratio is equalized across space, and we assume that this means that there is no incentive to migrate[5] since less expensive homes entail a corresponding reduction in disposable income. To find disposable income levels commensurate with spatial equilibrium, let us assume that if the house price (p_i) to real wage per worker ($\omega_i \kappa_i$) ratio is higher in location i than in k, then under spatial equilibrium there must be additional disposable income, such as from borrowing, pensions, investment income and suchlike, that allows i's ratio to exceed that of k. The ratio of i's price to real wage ratio to that of k provides the amount (π_i) by which we should in effect multiply i's wage rate to obtain i's

---

[4] According to the UK's Regional Accounts Methodology Guide, Gross Disposable Household Income (GDHI) is the amount of money that individuals (i.e. the household sector) have available for spending or saving. This is money left after expenditure associated with income, e.g. taxes and social contributions, and property ownership and provision. This income comes from both paid employment and through the ownership of assets or receipt of pensions and benefits. The largest single component of income received by the household sector in the UK in 2005 was compensation of employees, but this amounted to only 55% (UK National Accounts, Blue book, 2006).

[5] In contrast, in the new economic geography set out in FKV, the short –run equilibrium which produces spatial disparities in real wages, but only after labour migration in response to real wage differences do we reach a long-run equilibrium.

disposable income consistent with spatial equilibrium[6]. Assuming also that Y_i in each area depends on income additional to wages, and the number of varieties and hence G_i^M depends on the number of labour efficiency units $\kappa_r \lambda_r$ employed, the simultaneous equations are

$$\pi_{i,t} = \frac{p_{i,t}/(\omega_{i,t}\kappa_i)}{p_{k,t}/(\omega_{k,t}\kappa_k)} \tag{18}$$

$$Y_{i,t} = \kappa_i \lambda_i w_{i,t}^M \pi_{i,t} \tag{19}$$

$$w_{i,t}^M = [\sum_r Y_{r,t}(G_{r,t}^M)^{\sigma-1} T_{Mir}^{1-\sigma}]^{\frac{1}{\sigma}} \tag{20}$$

$$G_{i,t}^M = [\sum_r \kappa_r \lambda_r (w_{r,t}^M T_{Mir})^{1-\sigma}]^{\frac{1}{1-\sigma}} \tag{21}$$

$$\omega_{i,t} = w_{i,t}^M (G_{i,t}^M)^{-1} \tag{22}$$

Subscript t signifies iteration t, and the solution to (18) to (22) is said to occur when $Y_{i,t} \approx Y_{i,t-1}$, $w_{i,t}^M \approx w_{i,t-1}^M$, $G_{i,t}^M \approx G_{i,t-1}^M$ and $\pi_{i,t}^M \approx \pi_{i,t-1}^M$.

Application

The numerical solutions and simulations of shock effects on house prices are carried out using data for small administrative districts in England[7].

Measuring labour efficiency κ_r

The first consideration is relative labour efficiency, which is a set of fixed quantities in subsequent estimation. We assume that wage rates per efficiency worker are determined by both market potential and labour efficiency. In order to get a measure of

---

[6] In this, only i varies whereas k represents the City of London throughout.
[7] These are the Unitary Authority and Local Authority Districts, or UALADs, of which there are 353 in England.

relative labour efficiency *per se* in terms of relative wage rates, we eliminate the effect of market potential on wages and look at the adjusted wage rate ratios. To maintain simplicity the market potential measure adopted is

$$P_i = \sum_{r=1}^{R} w_r^o \lambda_r \exp(-\delta d_{ir})$$

$$d_{ir} = 0, i = r$$

(23)

in which w_r^o is observed wages, d_{ir} is the straight line distance in miles between (the centres of) i and r and $\delta = 0.05$ is a scalar with value chosen so that areas separated by 100 miles or more to have a minimal contribution to market potential. Approximation P will undoubtedly contain measurement error, and is by definition endogenous in the regression of P on observed wages w_r^o, and therefore we carry out 2sls estimation using a single instrument, equal to the area (measured in sq.km) of each UALAD.

We regress[8] log observed wages $\ln w^o$ on the fitted first stage values of log market potential[9] $\ln P$ and use the wage ratio

$$\exp(\hat{\varepsilon}) = \exp\left[\ln w^o - \left[\hat{b}_0 + \hat{b}_1 \ln P \right] \right]$$

(24)

to give relative labour efficiency via

$$\kappa_r = \frac{\exp(\varepsilon_r)}{\exp(\varepsilon_k)}$$

(25)

in which region k is a numeraire[10].

These κ values are illustrated in Figure 1, which suggests that labour is relatively more efficient in the South East of England, although evidently there are scattered pockets of 'high efficiency' elsewhere.

---

[8] For simplicity, we do not control explicitly for separate covariates representing the causes of labour efficiency variations.

[9] $\hat{b}_0 = 4.159$, $\hat{b}_1 = 0.083$, correlation observed and fitted values = 0.645

[10] The City of London.
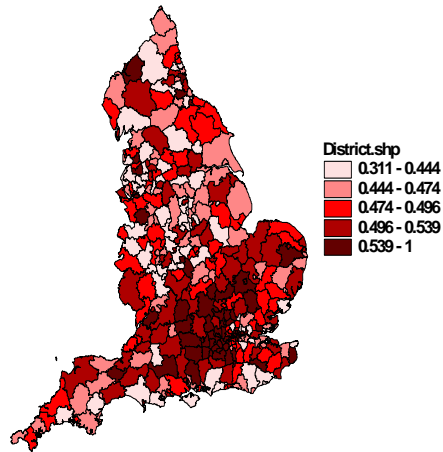
Figure 1 labour efficiency by UALAD

Iterative solution

Given an assumption of spatial equilibrium, it is assumed that the spatial distribution of labour efficiency units $(\kappa_r \lambda_r)$ is exogenous and fixed. Accordingly, and also given fixed numerical values for the exogenous terms T_{Mir} and σ, the solution to equations (18…22) is obtained iteratively since there is no obvious analytical solution. Iteration invariably solves the equations in the sense that after a number of iterations the equations produce endogenous outcomes with steady-state values, at which point they are terminated. Step 1 of the first round of iteration ($t = 1$) chooses initial wages per efficiency unit of labour $w_{i,0}^M$ equal to 1 (and likewise ω) thus allowing an initial estimate of house prices $p_{i,t}$ (see below) hence $\pi_{i,t}$ via equation (18). These then allow calculation of (19), namely $Y_{i,t}$. Step 3 uses these $Y_{i,t}$ values to obtain $w_{i,t}^M$ (20), using also T_{Mir} and σ and the price index $G_{i,0}^M$, which is an initial guess of $G = 1$ for all UALADs. The Fourth step of round 1 provides estimates of $G_{i,t}^M$ (21) using $w_{i,t}^M$ from (20). The fifth step calculates ω using $w_{i,t}^M$ and $G_{i,t}^M$ as in (22).

For the second round of iteration[11] ($t = 2$), given $w_{i,t-1}^M$ and the house price model coefficients and variables we obtain $p_{i,t}$ and use $\omega_{i,t-1}\kappa_i$ to give $\pi_{i,t}$. This then allows

---

[11] In each iteration, all variables are normalized by dividing by their value for the City of London.

(19),(20),(21) and (22) to be recalculated. In subsequent rounds of iteration ($t = 3,\ldots,T$) the same steps are applied, using the estimates of the preceding round for (18). The iterations terminate when the values of the endogenous variables $G_{i,t}^{M}$, $w_{i,t}^{M}$, $\pi_{i,t}$ and $Y_{i,t}$ reach steady state. The stopping criterion giving T is when $\sum_{i}(G_{it}^{M}-G_{it-1}^{M})^{2}<10^{-7}$,

$\sum_{i}(w_{it}^{M}-w_{it-1}^{M})^{2}<10^{-7}$, $\sum_{i}(\pi_{it}-\pi_{it-1})^{2}<10^{-7}$, and $\sum_{i}(Y_{it}-Y_{it-1})^{2}<10^{-7}$ simultaneously.
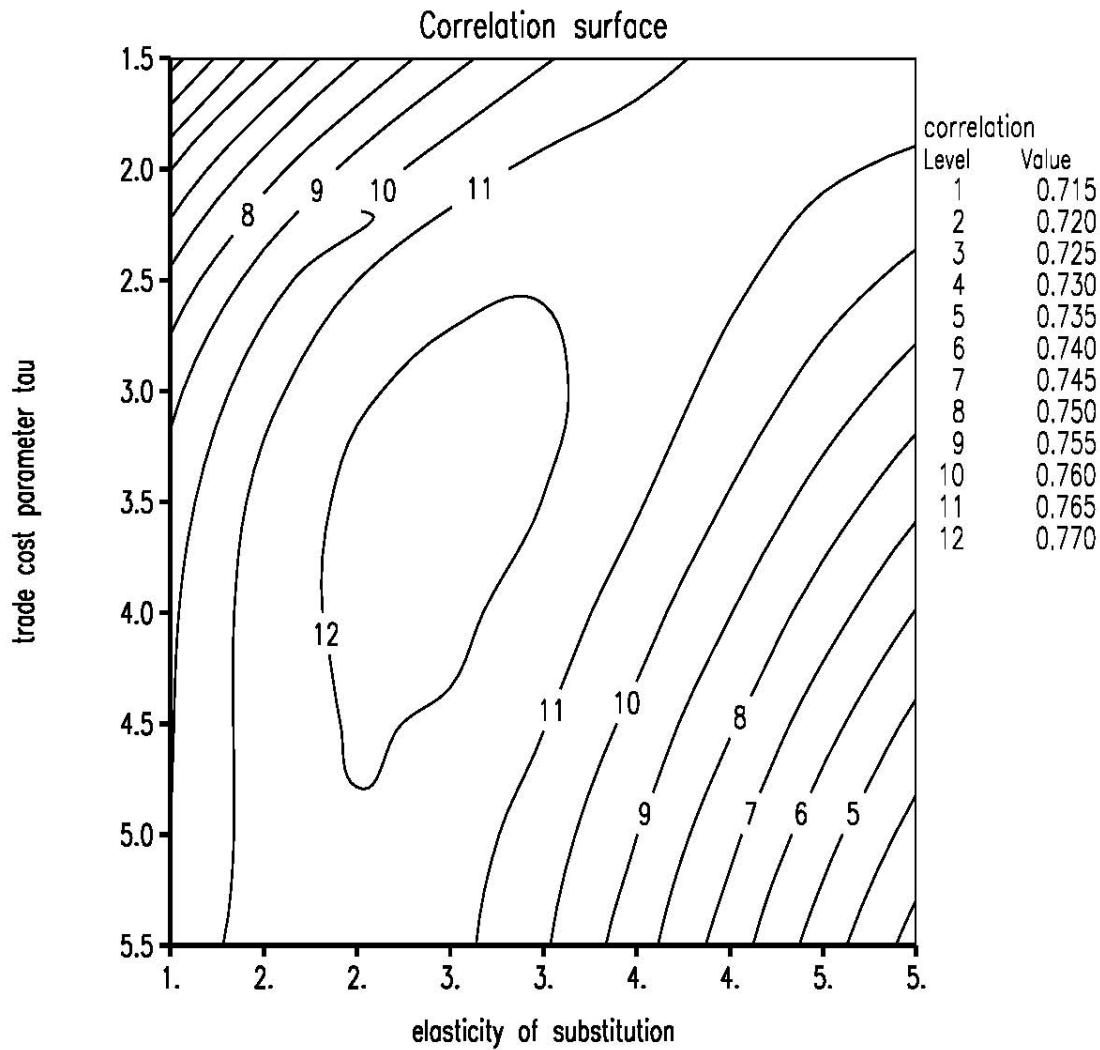
Figure 2 Correlation surface

Estimating the elasticity of substitution and trade cost function

In order to carry out the preceding solution leading to equilibrium, we need to know the values of the exogenous terms $T_{Mir} = \exp(\tau d_{ir})$ (trade costs) and σ (elasticity

of substitution). In the trade cost function, d_{ir} is the straight line distance measured in miles/1000. The optimal values of τ and σ are also obtained by an iterative search. For any single combination of σ and τ values, we solve equations (18…22) and then calculate the Pearson product moment correlation between $\kappa_r w_r^M$ and actual observed wage levels. Figure 2 shows the correlation surface, based on a 10 x 10 matrix of correlations obtained over a range of values of σ and τ, indicating that, approximately, $\sigma = 3.0$ and $\tau = 3.5$.

The house price model

a) specification

The specification is a simplified version of the house price model of Fingleton(2008), taking account also of the work of Cheshire and Sheppard (2004) and Gibbons and Machin(2003) among others. The model gives house prices[12] p_j on the left hand side via the reduced form from equilibrium housing supply and demand levels q_j. On the demand side, assume that q_j depends on income level Y_j^c which equals the sum of income (observed wage levels w_j^o times number of workers λ_j) within j and income weighted by commuting distance (D_{jk}) between employment in UALAD k and place of residence j, summing across all R UALAD's, so that

$$Y_j^c = \sum_{k=1}^{R} \exp(-\delta D_{jk}) w_k^o \lambda_k, D_{jk} = 0, j = k \qquad (26)$$

Assigning a value $\delta = 0.05$ has the effect of giving approximately zero weight beyond 100 miles. Note that with $D_{jk} = 0, j = k$, income is not down-weighted by within-area commuting distances. Assume also that demand for housing in j depends

---

[12] Used to calculate π_r (equations (18)).

12

negatively on j's price level (p_j) and is positively related to j's amenity (A_j). Other

unmodeled factors are represented by stochastic disturbances $\omega_1 \sim iid(0, \sigma_1^2 I)$. Assuming

that the relationship between prices and quantities is linear in natural logarithms[13], the

demand function is

$$\ln q_j = a_0 + a_1 Y_j^c + a_2 A_j - a_3 \ln p_j + \omega_1 \qquad (27)$$

The supply function

$$\ln q_j = b_0 + b_1 \ln p_j + b_2 O_j + \varsigma \qquad (28)$$

assumes that the level of housing supply increases with p_j and with the stock of

properties (O_j) and that other unmodeled effects are captured by the disturbance term

$\varsigma \sim iid(0, \sigma_2^2 I)$.

Normalizing the supply function with respect to p thus

$$\ln p_j = \frac{1}{b_1} \ln q_j - \frac{b_0}{b_1} - \frac{b_2}{b_1} O_j + \frac{\varsigma}{b_1} \qquad (29)$$

and substituting for $\ln q_j$ gives

$$\ln p_j = c_1 [a_0 + a_1 Y_j^c + a_2 A_j - a_3 \ln p_j + \omega_1] - c_0 - c_2 O_j + \xi$$

Simplifying this equation gives

$$\ln p_j = d_0 + d_1 Y_j^c + d_2 A_j - d_3 O_j + \varepsilon_j$$
$$\varepsilon_j \sim iid(0, \varpi^2 I) \qquad (30)$$

b)   Estimation

The price data[14] p_j are the average 2001 selling prices (all property types) by

UALAD[15], data which are provided by the UK's Land Registry. Income by area (Y_r) is

---

[13] This produces a better fit to the data that a linear relationship, gives a constant elasticity and avoids
negative prices.
[14] Data provided by the UK's Land Registry.
[15] 353 unitary authority and local authority districts.

taken to equal the local wage rate (w_r^o) times the local employment level (λ_r). The observed wages are taken from the 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 and employment levels are available on the NOMIS website (the Office for National Statistics' on-line labour market statistics database). The variable Y_r^c is normalised by dividing by the value for the City of London. The variable O which is equal to the number of owner-occupier households reported in the 1991 Census of Population[16]. The level of amenity A_j is given by three separate variables, 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), so that the estimated model becomes

$$\ln p_j = e_0 + e_1 Y_j^c + e_2 A_{Ej} + e_3 A_{Sj} + e_4 A_{Lj} - e_5 O_j + \varepsilon_j \tag{31}$$

---

[16] Local Base Statistics, Table L20 Tenure and amenities: Households with residents; residents in households. This is available in the NOMIS database.

c)   Results

Table 1 gives the OLS and 2sls estimates of equation (24).

Table 1. Initial Estimates of house price models

Dependent variable

	parameter est.	t ratio	parameter est.	t ratio
	OLS		2sls	
Constant	5.4175	10.70	5.4343	5.55
Y_j^C	0.0005957	13.77	0.0005847	12.97
O	-0.000001253	-2.14	-0.000001245	-2.10
A_E	1.6025	12.34	1.5993	6.36
A_S	5.1859	7.00	5.1570	6.72
A_L	-0.000004310	-13.39	-0.000004340	-13.07
$R^2, \overline{R}^2$	0.703		0.7070	
Standard Error	0.242		0.2403	
Log likelihood	2.4471		-------	
Residual correlation	I = 19.41		Z = 18.42	
Degrees of freedom	347		347	

Notes :

$\overline{R}^2$ = Squared Correlation actual and fitted. For the OLS model we use the conventional R^2 statistic.
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 an endogenous variable (no spatial lag).
The spatial autocorrelation tests use the matrix W defined in the Appendix.
The instruments used in the 2sls and for the test of 2sls residual spatial autocorrelation comprise the exogenous variables (constant,

O, A_S, 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).

Fitting the model by OLS does not take account of the possibly endogenous variables A_S and Y_j^C. The 4th and 5th columns of Table 2 give the two-stage least squares (2sls) estimates. The indication is that all the explanatory variables are significant and appropriately signed and that endogeneity is not a significant problem. However there is significant residual spatial autocorrelation, which may be a consequence of omitted spatially autocorrelated regressors, in which case the parameter estimates may be biased.

d)   Allowing spatial interaction

It is assumed that the significant residual spatial autocorrelation reported in Table 1 is a manifestation of demand and supply being displaced from where they would otherwise be. If prices 'nearby' , at k, are relatively high compared with j prices, this will push demand out from k into j . We refer to this as a displaced demand effect. We model this by assuming that j's demand is positively related to the weighted average of neighbours' prices, equal to the j'th cell of the vector $W_1 \ln p$, in which W_1 is a weighting matrix based on distance.

$$\ln q_j = a_0 + a_1 Y_j^c + a_2 A_j - a_3 \ln p_j + v W_1 \ln p_j + \omega \qquad (32)$$

Likewise, we envisage a displaced supply effect in which relatively high k prices will pull supply out from j in to k. In other words relatively high neighbours' prices $W_2 \ln p$ causes housing supply that would otherwise locate in j to locate instead in j's neighbours, thus giving a negative relationship between j's housing supply and the weighted average of neighbouring prices, hence

$$\ln q_j = b_0 + b_1 \ln p_j + b_2 O_j - \eta W_2 \ln p_j + \varsigma \qquad (33)$$

Normalizing the supply function with respect to p gives

$$\ln p_j = \frac{1}{b_1} \ln q_j - \frac{b_0}{b_1} - \frac{b_2}{b_1} O_j + \frac{\eta}{b_1} W_2 \ln p_j - \frac{\varsigma}{b_1} \qquad (34)$$

and substituting the quantity supplied by the quantity demanded q_j gives

$$\ln p_j = c_1[a_0 + a_1 Y_j^c + a_2 A_j - a_3 \ln p_j + v W_1 \ln p_j + \omega] - c_0 - c_2 O_j + c_3 W_2 \ln p_j - \xi \qquad (35)$$

16

in which A_j is a composite amenity variable. Simplifying by assuming that $W_1 = W_2 = W$, as defined in the Appendix, gives

$$\ln p_j = \rho \sum_{k \neq j} W_{jk} \ln p_k + d_0 + d_1 Y_j^c + d_2 A_j + d_3 O_j + \varepsilon_j \tag{36}$$

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

$$\ln p_j = \rho \sum_{k \neq j} W_{jk} \ln p_k + d_0 + d_1 Y_j^c + d_2 A_{Ej} + d_3 A_{Sj} + d_4 A_{Lj} + d_5 O_j + \varepsilon_j \tag{37}$$

Or equivalently in matrix terms we have

$$\ln p = (I - \rho W)^{-1}(Xd + \varepsilon) \tag{38}$$

in which X is an n by k matrix[17], I is the n by n identity matrix, d is a k by 1 vector of parameters, ρ is a scalar parameter and the disturbances $\varepsilon \sim iid(0, \tau^2 I)$ allow for measurement error in the price variable and for other unmodeled effects with variance τ^2.

---

[17] $k = 6$, $n = 353$.

Table 2. Estimates of house price models

Dependent variable	$\ln p$			
	parameter est.	t ratio	parameter est.	t ratio
	ML		2sls	
Constant	-1.9792	-3.75	-2.6976	-1.40
Y_j^c	0.0001948	5.20	0.0005752	14.65
O	0.000000071	0.17	-0.000000201	-0.36
A_E	1.4102	14.60	1.8457	8.20
A_S	3.5599	6.52	4.3583	6.33
A_L	-0.000001231	-4.20	-0.00000066	-0.79
ρ	0.6960	18.16	0.6014	4.70
R^2, $\bar{R}^2$	0.8413		0.8025	
Standard Error	0.1769		0.2108	
Log likelihood	96.5174		-------	
Residual correlation	LM = 1.336		Z = 1.371	
Degrees of freedom	346		346	

$\bar{R}^2$ = Squared Correlation actual and fitted. For the OLS model we use the conventional R^2 statistic.
LM is distributed as chi-squared 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.

For 2sls, the endogenous variables are Y_j^c, A_E and the spatial lag of house prices. The instruments are the exogenous variables, and 46 county dummies, and for the spatial lag we use the exogenous variables and their spatial lags obtained by multiplying by W.

Table 2 gives both ML[18] and 2sls estimates for the model, although because ML takes account only of the endogeneity of the spatial lag, attention is focussed on the 2sls

---

[18] Assuming normality for the iid disturbances.

estimates. These show that there is a significant spillover effect ($\rho \neq 0$), and that the coefficient signs are as anticipated, with negative values housing supply (*O*) and distance from London(A_L), although the 2sls estimates show that these variables are insignificant.

Eliminating these insignificant variables gives

$$\ln p_j = \rho \sum_{k \neq j} W_{jk} \ln p_k + d_0 + d_1 Y_j^c + d_2 A_{Ej} + d_3 A_{Sj} + \varepsilon_j$$

Table 3. Estimates of house price models

Dependent variable	$\ln p$			
	parameter est.	t ratio	parameter est.	t ratio
	ML		2sls	
Constant	-2.9000	-6.37	-4.0608	-4.52
Y_j^c	0.000201	5.21	0.0005769	14.31
A_E	1.4429	14.67	1.9029	8.69
A_S	2.6990	5.40	4.1849	6.54
ρ	0.7610	23.79	0.6967	15.47
R^2, $\bar{R}^2$	0.8353		0.7998	
Standard Error	0.1805		0.2185	
Log likelihood	85.4999		-------	
Residual correlation	LM = 1.807		Z = 1.166	
Degrees of freedom	346		346	

The estimates $\hat{\rho}, \hat{d}$ given in Table 3 are used to calculate $\pi_{i,t}$ (equation 18) via

$$\ln p_t = (I - \hat{\rho}W)^{-1}(X_t \hat{d}) \tag{39}$$

in which X_t denotes the value of matrix X for the t'th iteration, which changes because the second column is equal to

$$Y_{j,t}^c = \sum_{k=1}^{R} \exp(-\delta D_{jk})\kappa_k \lambda_k w_{k,t-1}^M \tag{40}$$

Model outcomes

The outcomes for the endogenous variables are presented as maps of the 353 UALADs. Figure 3 shows the distribution of 'market potential' (equal[19] to $\sum_r Y_{r,t}(G_{r,t}^M)^{\sigma-1}T_{Mir}^{1-\sigma}$, embodied in equation 20), taking account of labour efficiency units and disposable income. It highlights the concentration in and around London and the impact of inaccessibility to more peripheral UALADs. Figure 4 shows the income distribution (equation 19), reflecting the large concentration of workers in cities, wage levels and disposable income. Figure 5 shows the price index (equation 21), showing that it costs more to gain the same level of utility in the peripheral areas, because of the 'love of variety' which is more abundant in cities and their surrounds. Figure 6 gives the wage per efficiency unit, which is a direct function of market potential. Figure 7 is the observed wage level w_k^o, which can be compared with the nominal wage level (Figure 8) which is the endogenous model outcome equal to $\kappa_r w_r^M$. Figure 9 gives real wage per worker and Figure 10 is the house price given by $\exp(\ln p_T)$ where $\ln p_T = (I - \hat{\rho}W)^{-1}(X_T \hat{d})$. Figure 11 is the observed house price distribution. Figure 12 shows the (model-based) price to wage ratio ($\pi_{i,T}$) indicating those areas (mainly in the South East of England, but also in the rural South West and rural North) where under the equilibrium assumption, disposable income is evidently greater than indicated by wage rates. It also highlights those areas, particularly Northern industrial towns and remote parts of Eastern England, where under the equilibrium assumption there is a paucity of additional sources of disposable income and presumably high levels of debt.

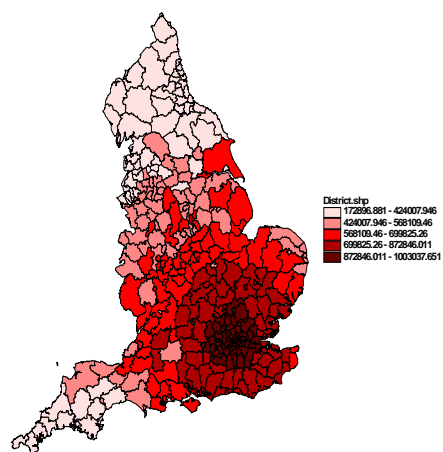[19] Relative to the City of London, then multiplied by 1000000.
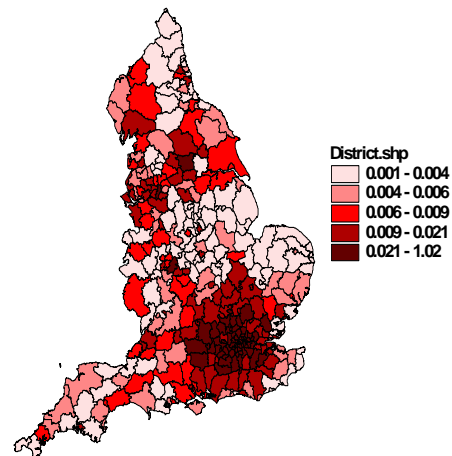
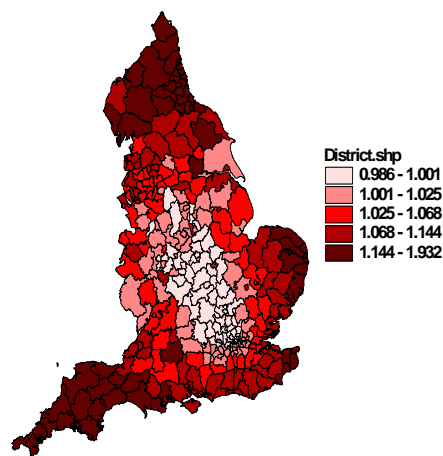Figure 3: Market potential

Figure 4 Income
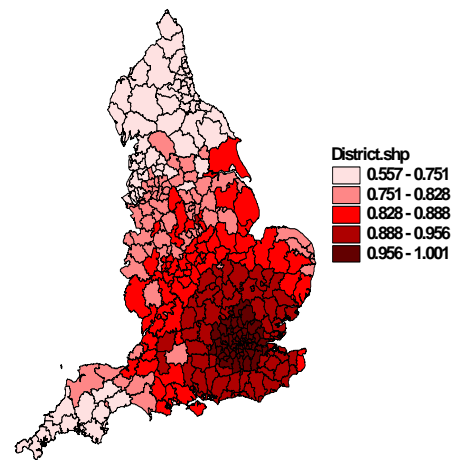

Figure 5 Price index
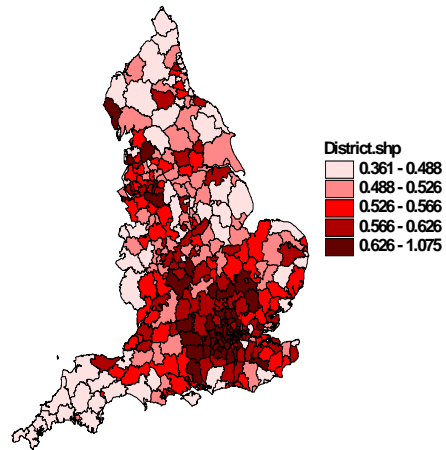

Figure 6 Wages per labour efficiency unit
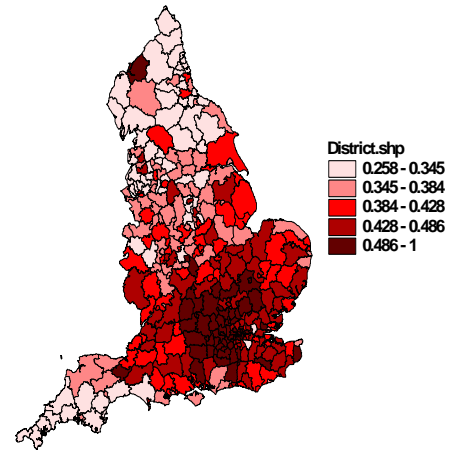
Figure 7 Observed wages
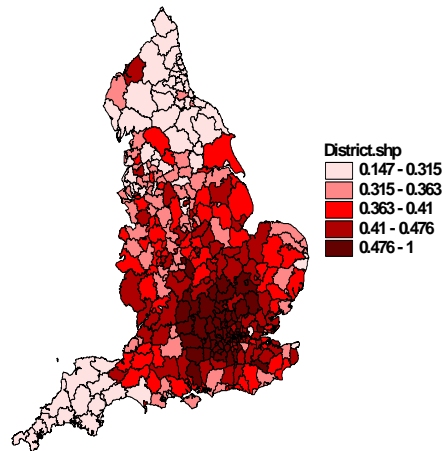

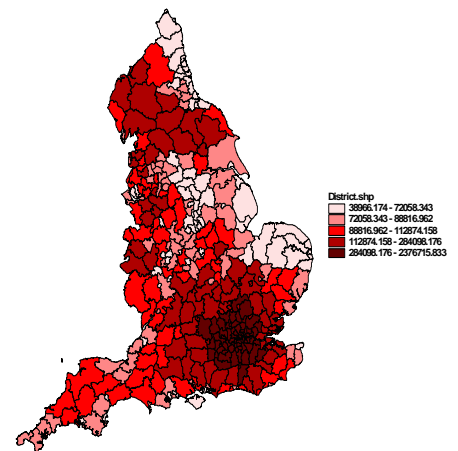Figure 8 Nominal wage per worker

Figure 9 Real wage per worker

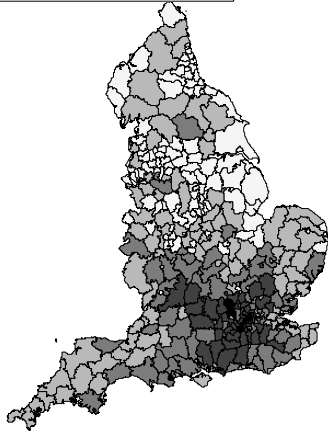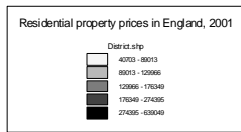Figure 10 House prices (model)
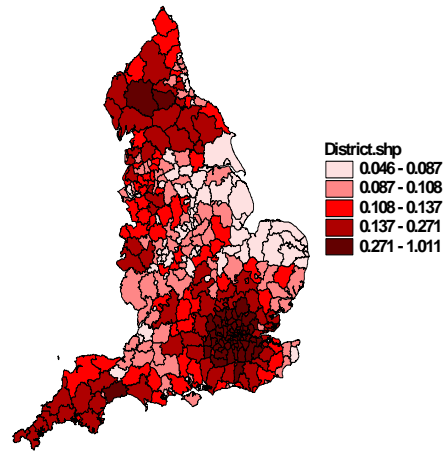
Figure 11 House prices   Figure 12 House prices relative to real wages per worker

Simulation

To illustrate an application of the model, assume that employment falls by 10% across all London UALADs, as shown by Figure 13. The model tells us that house prices fall over a much wider area of the South East of England and by as much as 26% in London (Figure 14) as a consequence of falling demand which is a combination of lower wages and lower employment (equation 40), and also because of the externalities causing house price changes to spill over to nearby UALADs. Wages depend on market potential (equation 20) which is also negatively affected by the negative shock to employment. Figure 15 shows the change in market potential relative to the City of London, which is why in relative terms market potential change is positive as one moves away from London. House prices in relation to real wages also more fall in and near London, giving the positive differences[20] shown by Figure 16.

---

[20] Figure 16 gives 'growth' in house prices in j minus 'growth' in house prices in City of London, minus change in the log of real wage per worker relative to City of London before and after the employment shock. So house prices in Luton fall by -10.2652% compared with -26.2728% for the City of London, and
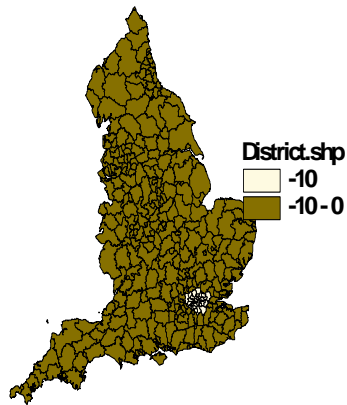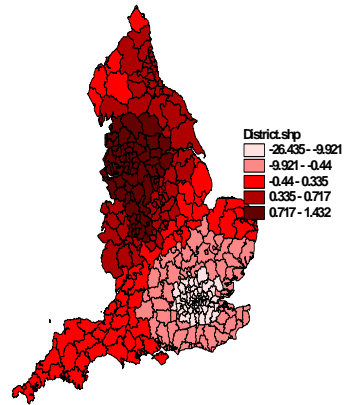
Figure 13 Employment change
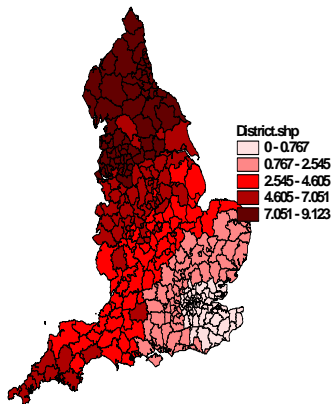

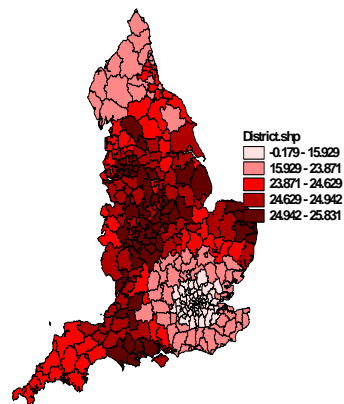Figure 14 House price change

Figure 15 Market potential (rel. ch.)

Figure 16 House price real wage ratio (rel. ch)

the change in relative real wage per worker is 0.5833% for Luton and 0 for the City of London, thus giving a difference of 15.4243%.

Conclusions

New economic geography theory is somewhat difficult to operationalize for various reasons. One is that there are several exogenous unknowns, such as the elasticity of substitution and the transport cost function. Another reason is the difficulty of deciding which sector is under a monopolistic competition market structure and which sector is competitive. While these operational decisions may be relatively easy in the international context, when modelling small local urban economies as in this paper, any decision seems somewhat more arbitrary. Also, at the level of cities, the price of property becomes a major aspect of economic decision making, and the role of agriculture is minimal, and while there is some literature which embodies the property sector within an NEG framework, it does seems lacking in terms of the causal variables that typical have been used to explain house price variation. Moreover, the basic theory as set out by FKV takes a very simple view of the causes of wage level differences, and the assumption that migration will lead to stable equilibria from a short-run equilibrium in which real wage differences exist does not seem to accord with the reality of the urban economy, which is for the UK at least essentially fairly static in terms of the long-run and stable patterns of wage and price inequality, that do not seem to be moving towards a long run equilibrium, but seem to be, approximately, in an equilibrium state already.

In this paper we endeavour to resolve some of the issues raised in attempting to operationalize the FVK model in several ways. The paper first of all takes the radical step of assuming that all firms are under monopolistic competition, with fixed costs and internal increasing returns to scale. This seems to be a realistic first approximation to the reality of the urban economy and has the benefit that with just one sector, the number of exogenous parameters is reduced, and therefore we can proceed to use a simple numerical technique to search for an optimum combination over a much reduced parameter space. Secondly, we use a very simple method to adjust for labour efficiency variations across space. Third, we do not treat the estimates obtained, with real wage differences across space, as a short run equilibrium, but as a stable equilibrium defined by the equality of the house price to disposable income ratio. Since house prices to real wages differ across space, under an assumption of equilibrium, we assume that this is because disposable

income differences exist and that these are quantifiable from the house price to real wage ratios.

The paper illustrates the equilibrium outcomes for the model using data for English local authority areas, and also gives the results of an small experiment in which employment in London receives a strong negative shock, falling by 10%. The resulting impact on house prices is strongly negative, not only in the immediate vicinity of the shock, but also further afield because of the externalities and spatial interactions embodied within the model.

.

References

Brakman, S, Garretsen, H, & M Schramm (2004) The Spatial Distribution Of Wages: Estimating The Helpman-Hanson Model For Germany, *Journal Of Regional Science*, 44 3 437–466

Cheshire P, Sheppard S (2004) Capitalising the Value of Free Schools: The Impact of Supply Constraints and Uncertainty, *Economic Journal* 114 499 F397-424

Fingleton, B (2008) Housing supply, housing demand, and affordability, *Urban Studies* 45 1545-1563

Gibbons S, Machin S (2003) Valuing English Primary Schools, *Journal of Urban Economics* 53 197-219

Glaeser E L (2008) Cities, *Agglomeration, and Spatial Equilibrium*, Oxford University Press, Oxford

Hanson, G H (2001) Scale Economies and the Geographic Concentration of Industry, *Journal of Economic Geography* 1 255–276

Hanson, G H (2005) Market potential, increasing returns and geographic concentration, *Journal of International Economics* 67 1 –24

Helpman, E (1998) The Size of Regions, in D. Pines, E. Sadka and I. Zilcha (eds.), *Topics in Public Economics*. Cambridge, U.K.: Cambridge University Press, pp. 33–54.

Appendix

A. Constructing the *W* matrix

The n by n matrix *W* is a row standardised version of matrix W^*, Hence for cell (*j,k*)

$$W_{jk} = \frac{W_{jk}^*}{\sum_k W_{jk}^*} \tag{2}$$

with $W_{jk}^* = \frac{1}{d_{jk}^2}$, 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 50km apart and falls quite sharply as distance increases.

B. Educational Attainment

This is based on the 1998 key stage 2 tests taken by 11-year-old pupils initially available for individual schools within smaller administrative areas nested within UALADs (these are known as wards, of which there are 8413 in England). The mean scores per Ward were then used to calculate mean scores for each of 353 English UALADs thus giving the regressor A_E.