

TESTING THE NEG MODEL : FURTHER EVIDENCE FROM PANEL DATA

Bernard FINGLETON *

***Abstract** - Local wage variations in the UK are explained by two non-nested rival hypotheses. The first derives from new economic geography theory, in which wages depend on market access. The second come from urban economics theory, giving a reduced form with wage rates dependent on employment density. The paper examines whether one of these rivals is encompassed by the other by fitting an artificial nesting model using three alternative panel data estimators. The estimates indicate that neither hypothesis is encompassed by its rival, suggesting a need for new, more comprehensive, theory.*

***Keywords** - PANEL DATA, SPATIALLY CORRELATED ERROR COMPONENTS, MARKET ACCESS, NEW ECONOMIC GEOGRAPHY, SPATIAL ECONOMETRICS, NON-NESTED HYPOTHESIS.*

***JEL Classification:** C33, O18, R15.*

* Department of Economics, University of Strathclyde (Scotland, UK). E-mail : bf100@cam.ac.uk

1. INTRODUCTION

Prior to the advent of the 'new economic geography' (NEG), spatial economic models allowing increasing returns to scale were somewhat detached from the main body of economic theory. More recently however, following a stream of research deriving initially from international trade theory and related sub-disciplines of economics, culminating in the publication of the book by Fujita, et al. (1999), it is fair to say that nowadays NEG has become widely recognized as a major contribution allowing, for the first time, increasing returns to scale to co-exist with micro-level theoretical assumptions. This has helped integrate spatial economics more fully into the main body of contemporary economic theory.

Recent research activity has also focused on operationalising, testing and refining NEG theory (see for example Combes and Overman, 2003; Head and Mayer, 2003, 2006; Redding and Venables, 2004; Rice and Venables, 2003; Brakman et al., 2006), and a central pillar of this work has been the so-called wage equation which, in the basic NEG model examined in this paper and described in Fujita et al. (1999), is one of a small number of simultaneous equations associated with the short-run equilibrium¹. The wage equation links nominal wages to market access or potential.

However, despite empirical evidence in support of the wage equation, this alone is an inadequate test of the explanatory power of NEG theory. A sharper test, following Head and Ries (2001), Davis and Weinstein (2003) and Fingleton (2006), is to examine the relative performance of NEG when confronted with a competing (and simpler) explanation of wage rate variations. In this spirit, the present paper tests two rival hypotheses, one associated with NEG's wage equation and the other originating from the urban economics literature (referred to here as the UE model). The UE model links wage rate variations to the density of employment, as a result of pecuniary externalities available in dense cities. There is a wide literature on the most appropriate way to test non-nested hypotheses. In this paper we use the simplest approach², described by Davidson and Mackinnon (2004) as 'inclusive regression', which simply puts the two competing non-nested hypotheses together in a single equation, in other words both market potential (NEG), and employment density (UE) together with covariates form part of an artificial nesting model (ANM).

The specific innovatory contribution of the paper is that the inclusive regression analysis is carried out using developments in panel data econometrics under parameter constraints consonant with commuting across local areas. Modelling panel data has several advantages, providing a richer information set than is available from purely cross-sectional analysis, and allowing inter-area (individual) heterogeneity to be modeled, thus avoiding potential omitted

¹ This occurs without any labour migration in response to real wage differentials. The long-run dynamics are driven by responses to real wage differences.

² Although the J test may be more powerful, inference is made more difficult by the non-standard distribution of the test statistic in finite samples.

variable bias. In the econometrics in the current paper, fixed effects (FE) and random effects (RE) panel models are fitted, together with a RE specification embodying spatially dependent error components (Kapoor et al., 2006; Fingleton, 2008) estimated by feasible generalized spatial two stage least squares (FGS2SLS) and generalized method of moments (GMM), although the same conclusions are arrived at under each of these three modelling approaches. Explicit consideration is given to simultaneous spillover effects across areas, which is an outcome of the small spatial units of observation³, so that inter-area commuting is recognized as a significant factor. The cost of allowing contemporaneous spillover is a reduced form ANM with parameter constraints. These are satisfied by using an iterative estimation approach, within which the three panel data estimators are embedded. Additionally, the FGS2SLS estimator allows for endogeneity and measurement error.

2. THE THEORETICAL MODELS

Both NEG and UE theory have a common basis in Dixit-Stiglitz monopolistic competition theory. They both assume a two-sector economy with a monopolistic competitive sector (M) comprising numerous small, differentiated firms with internal increasing returns to scale. This market structure assumption means that there is no strategic interaction, and free entry and exit, with competition driving profits down to zero in equilibrium. In the competitive (C) sector, we assume constant returns to scale and prices set on world markets.

2.1. The NEG model

The model implies that the level of M sector wages in area i , w_i^M , is related to i 's market potential P_i , which is one of the short-run⁴ equilibrium simultaneous equations (the wage equation) given by Fujita et al. (1999). In addition, M wages depend on A_i , which is area i 's level of efficiency, thus:

$$w_i^M = \left[\sum_r Y_r (G_r^M)^{\sigma-1} (\bar{T}_{ir})^{1-\sigma} \right]^{\frac{1}{\sigma}} = P_i^{\frac{1}{\sigma}} A_i \quad (1a)$$

$$\ln w_i^M = \frac{1}{\sigma} \ln P + \ln A \quad (1b)$$

As shown by Head and Mayer (2006), the presence of A in equation (1a) can be derived from micro-assumptions, by commencing with a labour quality adjusted production function for the firm. In (1), the summation is over areas

³ Comprising 408 unitary authority and local authority districts (UALADs), covering the surface area of Great Britain. There are 47 counties covering England, and 353 UALADs.

⁴ Fast entry and exit of firms drive profits to zero, but workers are slow to react and there is no migration in response to real wage differences between areas. Only in the long run would we expect movement to a stable long-run equilibrium with no real wage differences as a result of labour migration.

$r = 1, \dots, i, \dots, R$, the i to r transport cost is \bar{T}_{ir} , G_r^M denotes M prices, Y_r denotes income and σ is the elasticity of substitution for M varieties. In contrast, C goods are freely transported and produced under constant returns ; C wages w_i^C are constant across areas.

The price index is :

$$G_i^M = \left[\sum_r \lambda_r (w_r^M \bar{T}_{ir})^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (2)$$

with the number of varieties⁵ produced in area r denoted by λ_r . Income is denoted by:

$$Y_r = \theta \lambda_r w_r^M + (1-\theta) \phi_r w_r^C \quad (3)$$

in which ϕ_r is the share of C workers.

2.2. The UE model

UE theory explains productivity/wages variations by the varying supply of non-traded services to competitive industry. It therefore abstracts from transport costs, and consequently market potential/access is irrelevant. The core of UE theory is Cobb-Douglas production function:

$$Q = ((E^C A)^\beta I^{1-\beta})^\alpha L^{1-\alpha} \quad (4)$$

with Q denoting competitive industry production, $E^C A$ equal to C labour efficiency units, and I equal to the level of composite services. The quantity I is obtained via a CES production function for services under the monopolistic competition market structure. The parameter $\alpha < 1$ produces diminishing returns due to congestion effects (Ciccone and Hall, 1996), with variables measured per unit of land ($L^{1-\alpha} = 1$). Since I depends only on $E^M A$ and $N = A(E^C + E^M)$, then:

$$Q = ((E^C A)^\beta I^{1-\beta})^\alpha = \phi N^\gamma \quad (5)$$

with constants ϕ and $\gamma = \alpha[1 + (1-\beta)(\mu-1)]$, and with $\frac{\mu}{\mu-1}$ equal to the elasticity of substitution for different services. So long as $\gamma > 1$ there are increasing returns to employment density.

Using standard equilibrium theory, and with w^o denoting the overall wage rate:

⁵ Equal to the share in area r of the total supply of M workers.

$$\frac{w^o N}{Q} = \alpha \tag{6}$$

and taking natural logs, substituting for Q and⁶ $N = EA$ and simplifying gives:

$$\ln w_i^o = k_1 + (\gamma - 1) \ln E_i + (\gamma - 1) \ln A_i \tag{7}$$

in which k_1 denotes a constant.

3. LABOUR EFFICIENCY

The wage rate w_i^o refers to the wages paid by employers in area i , and these will therefore be influenced by workers commuting to work from outside i . Accordingly, labour efficiency within i depends on the characteristics of the workers resident in i , represented in equation (8) by X_{li}, \dots, X_{ki} , and also on the levels of labour efficiency in other areas from which workers commute.

In order to capture the contribution to efficiency from in-commuters, we allow efficiency in i to depend on efficiency in ‘neighbouring’ areas as follows:

$$\ln A_i = b_0 + b_1 X_{li} + \dots + b_k X_{ki} + \rho \sum_r \exp(-\delta_i D_{ir}) \ln A_r + \xi_i \tag{8}$$

$$r \neq i, D_{ir} \leq 100$$

in which we multiply by δ_i to allow for area i 's transport infrastructure. The values of these δ are determined by studying commuting flows, calibrating using census data, as in Fingleton(2003), and the weight $W_{ir} = \exp(-\delta_i D_{ir})$ is set to zero for $D_{ir} > 100$.

The adoption of a negative exponential function approximates to lower costs per mile for commuters at greater distance, possibly due to their willingness and ability to invest to achieve travel cost reductions, for example in season tickets available for long distance commuting that require an initial large outlay of money.

We express (8) more succinctly in matrix notation as:

$$\begin{aligned} \ln A &= \rho W \ln A + Xb + \xi \\ \xi &\sim N(0, \Omega^2) \\ (I - \rho W) \ln A &= Xb + \xi \\ \ln A &= (I - \rho W)^{-1} (Xb + \xi) \end{aligned} \tag{9}$$

⁶ E is the total employment level per square km.

An alternative to the endogenous spatial lag $W \ln A$ would be to use exogenous lags $(WX_{i_1}, \dots, WX_{i_k})$, but that excludes the unmodeled effects captured by ξ_i , and it is well known that for W with elements of which are less than 1 and $|\rho| < 1$ the spatial lag embodies the remote effects of exogenous variables going to infinity, since:

$$\ln A = (I - \rho W)^{-1} (Xb + \xi) = \left(\sum \rho^i W^i \right) (Xb + \xi) \quad (10)$$

The summation is from $i = 0$ to ∞ , $W^0 = I$, and in general W^i is the matrix product of W^{i-1} and W . We re-express this as :

$$\ln A = Xb + \rho WXb + \rho^2 W^2 Xb + \rho^3 W^3 Xb + \rho^4 W^4 Xb \dots \quad (11)$$

4. THE ARTIFICIAL NESTING MODEL (ANM)

The ANM combines the reduced forms from the NEG and UE models, giving:

$$\begin{cases} \ln w_{it}^o = k_2 + d_0 \ln E_{it} + d_1 \ln P_{it} + d_2 \ln A_{it} + \eta_{it} \\ \eta_{it} \sim N(0, \Gamma^2) \end{cases} \quad (12)$$

in which k_2 is a constant term, i refers to area and t refers to time. Substituting for $\ln A$ gives:

$$\ln w_{it}^o = k_2 + d_0 \ln E_{it} + d_1 \ln P_{it} + d_2 (I - \rho W)^{-1} (X_{it} b + \xi_{it}) + \eta_{it} \quad (13)$$

and multiplying by $(I - \rho W)$ obtains:

$$\begin{cases} (I - \rho W) \ln w_{it}^o = (I - \rho W) k_2 + (I - \rho W) (d_0 \ln E_{it} + d_1 \ln P_{it}) + X_{it} g + \zeta_{it} + (I - \rho W) \eta_{it} \\ \zeta_{it} \sim N(0, d_2^2 \Omega^2) \end{cases} \quad (14)$$

which can then be rearranged to give the ANM model:

$$\begin{aligned} \ln w_{it}^o = & (I - \rho W) k_2 + \rho W \ln w_{it}^o + d_0 [\ln E_{it} - \rho W \ln E_{it}] + d_1 [\ln P_{it} - \rho W \ln P_{it}] \\ & + X_{it} g + \zeta_{it} + (I - \rho W) \eta_{it} \end{aligned} \quad (15)$$

In (15), ρ determines the values of $(I - \rho W) k_2$ and the two compound variables $[\ln E_{it} - \rho W \ln E_{it}]$ and $[\ln P_{it} - \rho W \ln P_{it}]$, and defines the error process $(I - \rho W) \eta_{it}$ and the autoregressive interaction between $\ln w_{it}^o$ and

$W \ln w_i^o$. Thus in (15) we see the effect of the a priori specification (8), in that there are parameter constraints involving ρ .

5. DATA

In order to measure market potential, following equation (1a) we need transport cost $\bar{T}_{ir} = e^{\tau \ln D_{ir}} = D_{ir}^\tau$, M prices G_r^M , income Y_r and consequently wages w_i^M and w^C , the M and C worker shares λ_r and ϕ_i , the overall share of total employment engaged in M activities θ and also the elasticity of substitution for M varieties σ . We measure D_{ir} by straight line distance between areas, with $D_{ii} = \frac{2}{3} \sqrt{\frac{area_i}{\pi}}$, where $area_i$ is the number of square km in i . The use of the natural logarithm of distance gives a power function⁷, and this causes prices to increment by a factor equal to $\bar{T}_{ir} > 1$ provided $D_{ir} > 1$. This can always be achieved by a suitable choice of units (alternatively by using the function $\bar{T}_{ir} = 1 + D_{ir}^\tau$). We choose $\tau = 0.1$ partly on the basis that a distance of 100 miles will cause the delivered price to rise by a factor of 1.58. Larger values cause much greater increases and also increase the correlation between the resulting market potential and the competing UE variable. For example, choosing $\tau = 0.25$ is consonant with a factor of 3.16, which seems too large, and causes the correlation to increase from 0.6373 to 0.8328. A high correlation makes it more difficult to distinguish between the two rivals, and also means that market potential is dominated by internal demand, which is based on the arbitrary assumptions made about the value of D_{ii} .

Equation (3) shows that income depends on the number of varieties produced in area r , which is equal to the share in r of the total supply of M workers λ_r , and on the share of C workers ϕ_r together with the expenditure share of M goods (the overall share of total employment) θ . These quantities require definition of the M and C sectors. Accordingly, it is assumed that the M sector is equivalent to a subset of service sectors, while all other sectors are C activities. The subset is defined as the Banking, Finance and Insurance etc subgroup of the UK's 1992 Standard Industrial Classification. In contrast, in line with Fujita, Krugman and Venables (1999) the NEG literature commonly assumes M activities to be manufacturing, and all other sectors ('agriculture') to be C activities. However Redding and Venables (2004) use a composite of manufacturing and service activities, and in the UE literature (see for example Abdel-Rahman and Fujita, 1990) the M market structure assumptions are applied to services. Fingleton (2006) gives estimates that are robust to either definition of the M sector.

⁷ This introduces economies of distance and thus avoids some of the limitations of iceberg transport costs described by McCann (2005) and Fingleton and McCann (2007).

Equation (12) onwards proxies M wages w_i^M by w_i^o , as measured by the gross weekly pay (1999-2003) for male and female full time workers from UK's Office for National Statistics' New Earnings Survey. This is not sector specific, so we assume that the difference between w_i^o and w_i^M is measurement error captured by the error term. The wage rate for the C sector w^C , needed also to calculate Y_r , is constant across areas, is assumed to equal $MEAN(w_r^o)$.

It is assumed that the elasticity of substitution of M sector varieties is $\sigma = 6.25$, which is a central value among the range of estimates provided by the literature⁸ equal to the mid-point of the range given by Head and Mayer (2003). This is supported by the finding in Fingleton (2006) that the reciprocal of $\sigma = 6.25$ was within two standard errors of the estimated wage equation market potential coefficient. There are two main alternative approaches to obtaining the value of σ . One is direct nonlinear estimation, as carried out for example by Mion (2004) and Brakman et al. (2006), but this would be difficult to operationalize in the context of the iterative estimation methods to be described subsequently, and limit the potential complexity of the trade cost function that could be employed. The other alternative is the two-step linear estimation approach of Redding and Venables (2004), but this relies on bilateral trade flows which are unavailable at the level of spatial resolution being analyzed in this paper. Anderson and van Wincoop (2004) summarize various estimates, which are largely within the range 5 to 10. As with the assumed value $\tau = 0.1$, assuming $\sigma = 6.25$ does at any snapshot in time produce a market potential 'surface' that seems to accord reasonably well with a priori expectations, with a single peak centred on London, and a gradual incline to lower levels northwards and westwards. The central spine of Great Britain, from the South East to Greater Manchester, possesses the highest market potential levels. Assuming $\sigma = 20$ creates a surface with numerous sharp peaks centered on individual conurbations, which is too similar to the employment density surface to be able to properly differentiate the two. In contrast, assuming $\sigma = 2$ creates a more or less flat plain so that market potential has no explanatory power when entered into wage equation regressions.

With regard to the measurement of labour efficiency, within an area labour efficiency is assumed to depend on two key indicators. First, it depends on the level of formal education, since education provides skills which can be put to productive use. We use the log percentage of the population without any formal qualifications as an indicator, which is denoted by ln_ea , since this is an unambiguous measure available from the UK 2001 Census, and thus relatively free from measurement error or definitional problems. It is anticipated that ln_ea will be negatively related to wage rates. The disadvantage of this indicator is that it is constant across time, although even if data were available

⁸ Bröcker (2002) finds a value of 13, compared with 7.9 to 11.4 in Head and Ries (2001), 5 to 10 in Harrigan (1993), 9.28 in Eaton and Kortum (2002), 3 to 8.4 in Feenstra (1994), and 6.4 in Baier and Bergstrand (2001).

we would expect it to change very little over the period of the panel. The second component of worker efficiency relates to skills and knowledge gained at work. We focus on industries that are well known to be highly efficient and which provide transferable skills and efficiency generating capital equipment, particularly in information technology. Therefore the second element of the efficiency measure is the log of the location quotient for computing and related activities (1992 SIC 72) and research and development (1992 SIC 73), which is denoted by *lnknowlq*. The data come from the employee analysis in the UK's Annual Business Inquiry, available from NOMIS.

6. ESTIMATION

Three estimators are applied in this paper. One is the fixed effects (FE) estimator entailing individual (location) and time dummies (or equivalently mean deviations) and OLS. The second is the random effects (RE) estimator involving GLS (generalised least squares). The third is an estimator with random effects which allows for endogeneity and models spatially autocorrelated errors, involving feasible generalised spatial two stage least squares (FGS2SLS) and the generalised method of moments (GMM). In this section, we briefly sketch each estimation methodology, and give references to the relevant technical literature.

The FE estimation of the ANM model (equation (15)) involves some initial simplification, because k_2 is unknown and therefore the variable $(I - \rho W)k_2$ must be omitted⁹, therefore building spatial autocorrelation into the residuals. Also, the moving average error process $(I - \rho W)\eta$ is substituted by a simpler spatial independent normal error process. We re-introduce a spatially dependent error process subsequently. Consequently the FE specification is:

$$\begin{aligned} \ln w_{it}^o &= \rho W \ln w_{it}^o + d_0[\ln E_{it} - \rho W \ln E_{it}] + d_1[\ln P_{it} - \rho W \ln P_{it}] + \\ &\quad \alpha_i D_i + \phi_t T_t + X_{it} g + e_t \end{aligned} \tag{16}$$

$$e_t \sim N(0, \sigma^2 I)$$

in which w_{it}^o is the wage rate in area i at time t , $\alpha_i D_i$ is the fixed effect for i , with coefficients α_i and $N - 1$ place-dummy variables $D_i, i = 2, \dots, N$, and $\phi_t T_t$ is the fixed effect for time t , with coefficients ϕ_t and $T - 1$ time-dummies $T_t, t = 2, \dots, T$. The specification entails a constraint equalizing the coefficient ρ for $W \ln w_{it}^o, W \ln P_{it}$ and $W \ln E_{it}$. In order to satisfy this constraint, an iterative procedure is used¹⁰. The model is estimated by OLS.

Generalizing, we rewrite equation (16) as:

⁹ A simplification which is necessary in all the models estimated.

¹⁰ As for all models in this paper, using an approach described subsequently.

$$Y_t = \rho W Y_t + H_t \beta + e_t \quad (17)$$

in which $Y_t = \ln w_t^o$, $W Y_t = W \ln w_t^o$, H_t is the N by $(\tilde{k} = k + N + T)$ matrix of regressors $(\ln E_t - \rho W \ln E_t, \ln P_{it} - \rho W \ln P_{it}, X_{1t} \dots X_{kt}, D_2 \dots D_N, T_2 \dots T_N)$, β is a $(\tilde{k} \times 1)$ vector of parameters $(d_0, d_1, g_1 \dots g_K, \alpha_2 \dots \alpha_N, \phi_2 \dots \phi_T)$.

Since we are considering a panel with T periods rather than purely cross sectional data, we can omit t , hence:

$$\begin{aligned} Y &= \rho(I_T \otimes W)Y + H\beta + e = \tilde{X}b + e \\ \tilde{X} &= ((I_T \otimes W)Y, H) \\ b' &= (\rho, \beta') \end{aligned} \quad (18)$$

in which Y is a $(TN \times 1)$ vector of observations obtained by stacking $Y_t = \ln w_t^o$ for $t = 1 \dots T$, \tilde{X} is a $(TN \times (1 + \tilde{k}))$ matrix of regressors, comprising the $(TN \times 1)$ vector $(I_T \otimes W)Y$, and H which is a $(TN \times \tilde{k})$ matrix of regressors. Also b is the $((\tilde{k} + 1) \times 1)$ vector of parameters $(\rho, d_0, d_1, g_1 \dots g_K, \alpha_2 \dots \alpha_N, \phi_2 \dots \phi_T)$. In addition, given that I_T is a $(T \times T)$ diagonal matrix with 1s on the main diagonal and zeros elsewhere, and I_N is a similar $(N \times N)$ diagonal matrix, then $I_{TN} = I_T \otimes I_N$ is a $(TN \times TN)$ diagonal matrix with 1s on the main diagonal and zeros elsewhere.

We also estimate equation (16) with $e_t \sim N(0, \sigma^2 I)$ replaced by an autoregressive (AR) error process using ML (Baltagi, 2001, Elhorst 2003). In this specification, in each period $e_t = \lambda W_c e_t + \xi_t$, in which λ is an unknown parameter, and ξ_t is an $(N$ by $1)$ vector of time t innovations distributed as $N(0, \sigma^2 I)$ and W_c is a standardized contiguity matrix, as distinct from the ‘commuting W matrix’ described earlier. However the error process does not correspond to the moving average (MA) process of (15) and the estimation takes no account of the endogeneity of right hand side variables.

The fixed-effects estimator has the important advantage of allowing endogeneity of the regressors with respect to the individual effects. However it is limited by not being able to identify time-invariant regressors, since time-invariant regressors are aliased by fixed effect dummies. Also in this case there are 408 areas and therefore we need (in effect) 408 dummies (or equivalently deviations from the group means) to control for inter-area heterogeneity. There is a loss of a large number of degrees of freedom due to the fixed effects. This reduces the efficiency with which we can estimate the regression coefficients.

The RE estimator allows identification of time-invariant regressor parameters, and estimation of time-varying regressor parameters is carried out more efficiently, but exogeneity of the regressors with respect to the random individual effects is assumed to maintain consistency. The components of the errors ξ are:

$$\mu \sim iid(0, \sigma_\mu^2) \tag{19}$$

$$v \sim iid(0, \sigma_v^2)$$

$$\xi = (I_T \otimes I_N)\mu + v \tag{20}$$

in which μ is the $(N \times 1)$ vector of individual effects, v is the $(NT \times 1)$ vector of transient errors, I_T is a $(T \times 1)$ matrix with 1s, and $I_T \otimes I_N$ is a $(TN \times N)$ matrix equal to T stacked I_N matrices. The result is that the $(TN \times TN)$ innovations variance-covariance matrix Ω_ξ is non-spherical. The composite disturbance term means that OLS is not appropriate. We therefore use generalised least squares (GLS), specifically the Swamy and Arora (1972) estimator.

The third estimator (FGS2SLS plus GMM) allows spatial dependence in the errors and controls for endogeneity. For the error or disturbance process, one of two error process assumptions¹¹ made is that in each period $e_t = \lambda W_c e_t + \xi_t$, in which λ is an unknown parameter, and ξ_t is an $(N \times 1)$ vector of time t innovations. These create the $(NT \times 1)$ vector e :

$$e = (I_{TN} - \lambda I_T \otimes W_c)^{-1} \xi \tag{21}$$

in which ξ is an $(NT \times 1)$ vector of innovations.

Notice that the two error components μ and v are assumed to follow the same autoregressive process (cf. Baltagi and Li, 2006), since:

$$e = (I_{TN} - \lambda I_T \otimes W_c)^{-1} \xi = (I_{TN} - \lambda I_T \otimes W_c)^{-1} ((I_T \otimes I_N)\mu) + (I_{TN} - \lambda I_T \otimes W_c)^{-1} v \tag{22}$$

In estimating $\rho, \beta, \sigma_\mu^2, \sigma_v^2$ and λ , the method of Kapoor et al. (2007) involving GMM, nonlinear least squares and FGLS, is adapted to allow for endogenous variables in the matrix of regressors \tilde{X} (see Fingleton, 2008). Regressor matrix \tilde{X} includes vector $(I_T \otimes W)Y$ which is endogenous by definition, and the endogenous¹² variables $\ln E_t$ and $\ln P_t$ stacked for $t = 1 \dots T$.

¹¹ The autoregressive process (AR) and the moving average (MA) error process.

¹² Employment hence employment density will probably increase in response to high wages. Market access is dependent on wages by definition.

Hence we proceed using instrumental variables, but also take account of the non-sphericity of variance-covariance matrix Ω_ξ . To filter out the AR error dependence, we use a Cochrane-Orcutt (C-O) transformation, premultiplying by $I_{TN} - \lambda I_T \otimes W_c$ since $e = (I_{TN} - \lambda I_T \otimes W_c)^{-1} \xi$, therefore:

$$\begin{aligned} Y^* &= (I_T \otimes (I_N - \hat{\lambda} W_c)) Y \\ X^* &= (I_T \otimes (I_N - \hat{\lambda} W_c)) \tilde{X} \\ \xi &= (I_T \otimes (I_N - \hat{\lambda} W_c)) e \end{aligned} \quad (23)$$

A linearly independent subset of the exogenous variables is used to give the $(TN \times f)$, matrix of instruments Z , and we assume matrices \tilde{X} and Z are full column rank with $f \geq (\tilde{k} + 1)$. Alternatively, assuming a moving average (MA) error process, we pre-multiply by $(I_{TN} - \lambda I_T \otimes W_c)^{-1}$. Both of these provide $\hat{\beta}$ and hence \hat{e} , which forms part of the GMM estimating equations for λ , σ_v^2 , and σ_1^2 given in Kapoor et al. (2007).

Fingleton (2008) gives the GMM panel data estimator for MA errors with extension to an endogenous spatial lag.

7. RESULTS

Table 1 gives the OLS and ML estimates for the FE specification, with heterogeneity controlled by individual and time dummies. In addition, we give robust HAC standard errors due to Arellano (2003), although with N (the number of areas) large relative to T (the number of time periods), the distribution theory relies on cross-sectional independence. The Beck-Katz (1995) standard error estimates do allow for contemporaneous correlation across areas and for heteroscedasticity by area, assuming no time series autocorrelation, and in this case significantly reduce standard errors and increase t-ratios. However these will be relatively inaccurate for small T (here equal to 6) since a typical element of the contemporaneous covariance matrix is

estimated by $\hat{\Sigma}_{ij} = \sum_{t=1}^T e_{it} e_{jt} / T$. However each of these estimators indicates

that the rival compound variables $\ln E - \rho W \ln E$ (employment density) and $\ln P - \rho W \ln P$ (market potential) are significant, suggesting that neither of the competing hypotheses is dominant. The Wald test for joint significance of the time-dummies returns a test statistic equal to 14.7708, the test statistic is asymptotically distributed as χ_4^2 under the null of no significant time-dummy, and this null is rejected given the p-value equal to 0.005. The individual effects are also jointly significant, as the test statistic equals 13.139, with a near zero p-value in the $F(407, 1624)$ distribution, thus rejecting the null of a common intercept. The presence of the fixed individual effects means that the time

constant variable $\ln ea$ is aliased. Also the effect of $\ln knowlq$ is not significantly different from 0. The implication of the OLS point estimates is that doubling employment density causes wages to rise by more than 4%, since $\ln(2^{0.0626781}) = 0.043445$, and doubling market potential causes wages to increase by about 11.3%. However none of these point or standard error estimates make allowance for the potential endogeneity of variables WY , $\ln E$, and $\ln P$. For example, Beck-Katz (1995) standard errors assume consistent OLS for the residuals.

Table 1: Fixed effects (time-dummies and place-dummies) with iteration

VARIABLE	COEFFICIENT	STDERROR	T STAT	P-VALUE
WY	0.0583853 (0.066490)	0.0101635 (0.0086418) {0.0115423} [0.00556337]	5.745 (7.694) {5.058} [10.495]	<0.00001 *** (<0.00001 ***) {<0.00001 ***} [<0.00001 ***]
$\ln knowlq$	0.00244665 (0.001757)	0.00423387 (0.00373829) {0.00487733} [0.00453254]	0.578 (0.470) {0.502} [0.540]	0.56342 (0.6384) {0.61599} [0.58941]
$\ln E - \rho W \ln E$	0.0626781 (0.063334)	0.0240388 (0.0212245) {0.0301758} [0.0151206]	2.607 (2.984) {2.077} [4.145]	0.00921 *** (<0.003 ***) {0.03795 **} [0.00004 ***]
$\ln P - \rho W \ln P$	0.162578 (0.163211)	0.0286024 (0.0213739) {0.0332229} [0.0154805]	5.683 (7.636) {4.894} [10.502]	<0.00001 *** (<0.00001 ***) {<0.00001 ***} [<0.00001 ***]
λ	(-0.067970)		(-2.185)	(0.029)*
Rsq (squared correlation between fitted and actual)	0.9426525 (0.942612)			
Sum of squared residuals	3.20721 (3.20948)			

Notes: Estimates for fixed effects are omitted because of limited interest and to conserve space. ML estimates are given in ().

Robust HAC estimator given by Arellano(2003) in { }.

Standard errors using the estimator proposed by Beck and Katz (1995), see also Greene (2003, chapter 13) in [].

The ML estimates for the FE model with an AR error process, using a standardized contiguity matrix, give t-ratios that are larger than under HAC estimation, but mainly smaller than with Beck-Katz (1995) estimates. Compared with OLS, there is only a very minor impact on the point estimates, and we come to a similar conclusion that both UE and NEG variables are significant, although the presence of negative error dependence was not anticipated.

Table 2 gives the RE estimates, whereby individual heterogeneity is modeled explicitly as an error component. Given the presence of the UE (employment density) variable, the coefficient on market potential is clearly significant. All the other variables are significant, with appropriate signs. The indication is that doubling employment density causes wages to rise by less than

1%, as indicated by the fact that $\ln(2^{0.0137671}) = 0.0095$. Doubling market potential causes wages to increase by $\ln(2^{0.0805990}) = 0.05586$ or about 5.6%. However, as with the FE estimates of Table 1, these RE estimates do not take account of the endogeneity of variables, moreover there is no control for spatial dependence in the residuals. Also RE estimation imposes an assumption that the regressors are orthogonal to the random individual effects.

Table 2 : Random effects with iteration; time-dummies

VARIABLE	COEFFICIENT	STDERROR	T STAT	P-VALUE
<i>constant</i>	3.33823	1.02760	3.249	0.00118 ***
<i>WY</i>	0.00194853	0.000299869	6.498	<0.00001 ***
<i>ln_ea</i>	-0.155550	0.0194621	-7.992	<0.00001 ***
<i>lnknowlq</i>	0.0219163	0.00358608	6.111	<0.00001 ***
<i>lnE-ρWlnE</i>	0.0137671	0.00298089	4.618	<0.00001 ***
<i>lnP-ρWlnP</i>	0.0805990	0.0275316	2.928	0.00345 ***
Sum of squared residuals	14.4433			

Notes: Estimates for time-dummies omitted because of limited interest and to conserve space.

Table 3 : FGS2SLS and GMM estimates with iteration

VARIABLE	COEFFICIENT	STDERROR	T STAT	P-VALUE
<i>constant</i>	1.87711 (1.85422)	2.10043 (2.1219)	0.893677 (0.873849)	0.3715 (0.3822)
<i>WY</i>	0.00220046 (0.002445)	0.000570329 (0.000662915)	3.85823 (3.68825)	<0.0002 *** (<0.00025 ***)
<i>ln_ea</i>	-0.113757 (-0.11673)	0.0272751 (0.028223)	-4.17073 (-4.13601)	<0.00004 *** (<0.00004 ***)
<i>lnknowlq</i>	0.0495796 (0.0493868)	0.00569766 (0.00580332)	8.70176 (8.51009)	<0.00001 *** (<0.00001 ***)
<i>lnE-ρWlnE</i>	0.0132398 (0.0133824)	0.00352675 (0.00360731)	3.75411 (3.70981)	<0.0002 *** (<0.0002 ***)
<i>lnP-ρWlnP</i>	0.118037 (0.118839)	0.0559747 (0.0565922)	2.10876 (2.09991)	0.035 * (0.036 *)
λ	0.0294994 (-0.0303594)			
σ_v^2	0.0015 (0.00255578)			
σ_1^2	0.0238 (0.0246172)			
Square of correlation between fitted and actual	0.7412 (0.743854)			
Sum of squared residuals	14.6134 (14.3733)			

Notes: Estimates for time-dummies omitted because of limited interest and to conserve space. Estimates with MA errors given in ().

Table 3 summarizes the third estimator, namely FGS2SLS plus GMM. The specification allows for individual heterogeneity via random effects and time heterogeneity via dummies, Compared with the FE and RE estimators, it controls for the endogeneity of WY, lnE and lnP . Spatial error dependence is

modeled by AR and MA error processes. The table shows that both market potential and employment density are significant. In this case doubling employment density causes wages to rise by less than 1% and doubling market potential causes wages to increase by about 8.2%.

The instruments are based on the employment densities and market potentials for the year 1998, which precedes the estimation period and avoids reverse causality. As an added insurance against simultaneity and measurement error the 1998 values are coded¹³ 1, 0 and -1 according to whether they are above, between or below the upper and lower quartiles. This provides the instruments $g3le_$ and $g3lnMP_$. Additional instruments are created by premultiplying by a standardized contiguity matrix, thus giving $Wg3le_$ and $Wg3lnMP_$. Table 4 provides an illustration. Applying the Sargan over-identification test to the (iterative) 2SLS estimates indicates that, with a test statistic equal to 2.50 and p-value of 0.11 in the χ^2_1 distribution, these instruments are independent of the 2SLS residuals.

Table 4 : Instrument construction

$lnE(1998)$	$lnP(1998)$	$g3 le_$	$Wc g3 le_$	$g3lnMP_$	$Wc g3lnMP_$
7.8135	37.2766	1	-0.25	1	0.25
4.3853	37.2472	-1	-0.125	0	0.625
3.7853	37.0720	-1	-0.4286	-1	-0.4286
7.3780	37.0987	1	-1	-1	-1
5.8159	37.0871	0	-1	-1	-0.3333

Key :

$lnE(1998)$ = log employment density 1998

$lnP(1998)$ = log of market potential 1998

$g3 le_$ = -1 if below lower quartile, 0, 1 if above upper quartile of $lnE(1998)$

$Wc g3 le_$ = weighted average of $g3 le_$ in contiguous areas

$g3lnMP_$ = -1 if below lower quartile, 0, 1 if above upper quartile of $lnP(1998)$

$Wc g3lnMP_$ = weighted average of $g3lnMP_$ in contiguous areas

8. CONCLUSION

The evidence presented suggests that neither the UE nor the NEG model alone provides a satisfactory explanation of local wage variations in the UK. Each of the two rival hypotheses evidently carries additional explanatory information, with neither encompassing the other. Also it is evident that variations in labour efficiency (mediated by commuting) carry some independent explanatory power. These inferences are based on the outcome of three estimators, FE, RE and FGS2SLS plus GMM, each of which involves different assumptions. In the case of the FE and RE estimators, it is assumed that market potential and employment density are exogenous. Also, under the RE estimator, there is an assumption of exogeneity of the regressors with respect to the individual random effects and no allowance for spatial error dependence. Under FGS2SLS, we assume that finite sample bias is negligible

¹³ As discussed by Kennedy (2003).

under the specific instruments that happen to have been chosen to allow consistent estimation. Despite the caution these and other assumptions engender, the broad conclusions deriving from the various results obtained do not depend on the econometric model used, with each of these different estimators indicating some empirical advantage in hybrid models and the need for new, more comprehensive, theory.

REFERENCES

- Abdel-Rahman H., Fujita M., 1990, "Product variety, Marshallian externalities and city size", *Journal of Regional Science*, vol. 30, pp. 165-183.
- Anderson J.E., van Wincoop E., 2004, "Trade costs", *Journal of Economic Literature*, vol. 42, pp. 691-751.
- Arellano M., 2003, *Panel Data Econometrics*, Oxford University Press, Oxford.
- Baltagi B.H., 2001, *Econometric Analysis of Panel Data*, second ed., John Wiley New York.
- Baltagi B.H., Li D., 2006, "Prediction in the panel data model with spatial correlation: The case of liquor", *Spatial Economic Analysis*, vol., pp. 175-185.
- Baier S.L., Bergstrand J.H., 2001, "The growth of world trade: tariffs transport costs and income similarity", *Journal of International Economics*, vol. 53, pp. 1-27.
- Beck N., Katz J.N., 1995, "What to do (and not to do) with time-series cross-section data", *The American Political Science Review*, vol. 89, pp. 634-647.
- Brakman S., Garretsen H., Schramm M., 2006, "Putting New Economic Geography to the test", *Regional Science and Urban Economics*, vol. 36, pp. 613-635.
- Bröcker J., 2002, Spatial effects of European transport policy: A CGE Approach, dans Hewings G.J.D., Sonis M., Boyce D. (eds.), *Trade Networks and Hierarchies - Modelling Regional and Interregional Economies*, Springer, Berlin, pp. 11-28.
- Ciccone A., Hall R.E., 1996, "Productivity and the density of economic activity", *American Economic Review*, vol. 86, pp. 54-70.
- Combes P.P., Overman H., 2003, The spatial distribution of Economic activity in the EU, Discussion Paper 3999, Centre For Economic Policy Research, London.
- Davidson R., Mackinnon J.G., 2004, *Econometric Theory and Methods*, Oxford University Press, Oxford.

- Davis D.R., Weinstein D.E., 2003, "Market access economic geography and comparative advantage: an empirical test", *Journal of International Economics*, vol. 59, pp. 1-23.
- Eaton B., Kortum S., 2002, "Technology, geography and trade", *Econometrica*, vol. 70, pp. 1741-1779.
- Elhorst P., 2003, "Specification and estimation of spatial panel data models", *International Regional Science Review*, vol. 26, pp. 244-268.
- Feenstra R.C., 1994, "New product varieties and the measurement of international prices", *American Economic Review*, vol. 84, pp. 157-177.
- 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*, vol. 58, pp. 501-530.
- Fingleton B., 2008, "A generalized method of moments estimator for a spatial panel model with an endogenous spatial lag and spatial moving average errors", *Spatial Economic Analysis*, vol. 3, pp. 27-44.
- Fingleton B., McCann P., 2007, "Sinking the iceberg? On the treatment of transport costs in new economic geography", dans Fingleton B. (ed.), *New Directions in Economic Geography*, Edward Elgar, Cheltenham, pp. 168-203
- Fujita M., Krugman P., Venables, A.J., 1999, *The Spatial Economy : Cities Regions and International Trade*, MIT Press, Cambridge, Mass.
- Greene W.H., 2003, *Econometric Analysis*, fifth ed., Prentice-Hall, USA, NJ.
- Harrigan J., 1993, "OECD Imports and trade barriers in 1983", *Journal of International Economics*, vol. 34, pp. 91-111.
- Head K., Mayer T., 2003, "The empirics of agglomeration and trade", Discussion Paper 3985, Centre For Economic Policy Research, London.
- Head K., Mayer T., 2006, "Regional wage and employment responses to market potential in the EU", *Regional Science and Urban Economics*, vol. 36, pp. 573-594.
- Head K., Ries J., 2001, "Increasing returns versus national product differentiation as an explanation for the pattern of US-Canada trade", *American Economic Review*, vol. 91, pp. 858-876.
- Kapoor M., Kelejian H.H., Prucha I., 2007, "Panel data models with spatially correlated error components", *Journal of Econometrics*, vol. 140, pp. 97-130.
- Kennedy P., 2003, *A Guide to Econometrics*, fifth ed., Blackwell, Oxford.
- McCann P., 2005, "Transport costs and new economic geography", *Journal of Economic Geography*, vol. 5, pp. 305-318.

- Mion G., 2004, "Spatial externalities and empirical analysis: the case of Italy", *Journal of Urban Economics*, vol. 56, pp. 97-118.
- Redding S., Venables A.J., 2004, "Economic geography and international inequality", *Journal of International Economics*, vol. 62, pp. 53-82.
- Rice P., Venables A. J., 2003, "Equilibrium regional disparities: theory and British evidence", *Regional Studies*, vol. 37, pp. 675-686.
- Swamy P.A.V.B., Arora S.S., 1972, "The exact finite sample properties of the estimators of coefficients in the error components regression models", *Econometrica*, vol. 40, pp. 261-275.

LE TEST DU MODÈLE DE LA NOUVELLE ÉCONOMIE GÉOGRAPHIQUE : RÉSULTATS SUPPLÉMENTAIRES SUR DONNÉES DE PANEL

Résumé – Les variations locales de salaire en Grande-Bretagne sont expliquées à l'aide de deux hypothèses rivales non emboîtées. La première dérive des modèles d'économie géographique, dans lesquels les salaires dépendent du potentiel de marché. La seconde provient des modèles d'économie urbaine, qui induisent une forme réduite dans laquelle les taux de salaire dépendent de la densité d'emploi. Cet article examine si l'une de ces théories rivales est incluse dans l'autre en estimant un modèle, incorporant ces deux hypothèses, à l'aide de trois estimateurs en données de panel. Les résultats indiquent qu'aucune hypothèse n'est englobée par sa rivale, ce qui suggère la nécessité d'une nouvelle théorie plus générale.