THE WAGE CURVE RECONSIDERED: IS IT TRULY AN 'EMPIRICAL LAW OF ECONOMICS'?

Bernard FINGLETON[*], Silvia PALOMBI [**]

Abstract - *The negative relationship between wages and unemployment, embodied within the so-called wage curve, has an extensive literature and has been referred to as 'an empirical law of economics'. However there are newer alternative theories that seek to explain the variation in wages without reference to unemployment rates, namely Urban Economics (UE) and New Economic Geography (NEG). The present paper considers the relative success of non-nested models of wage determination from the perspective of Great Britain's 408 local authorities over the period 1998-2010.*

Key-words: WAGE CURVE, EMPLOYMENT DENSITY, MARKET POTENTIAL, SPATIAL PANEL DATA ECONOMETRICS, NON-NESTED MODELS, SPATIAL J-TEST

JEL Classification: C23, C52, J30, J40, J63, O18, R11, R12

[*] Department of Land Economy, University of Cambridge, 19 Silver Street, Cambridge CB3 9EP, UK; bf100@cam.ac.uk

[**] Department of Economics, University of Strathclyde, Sir William Duncan Building, 130 Rottenrow, Glasgow G4 0GE, UK; silvia.palombi@strath.ac.uk

1. INTRODUCTION

Since Blanchflower and Oswald's seminal paper on *The Wage Curve* (Blanchflower and Oswald, 1990), a wealth of international research has emerged on the responsiveness of wages to changes in local labour market conditions. In most cases, an unemployment elasticity of pay of around -0.10 has been found; given its uniformity across countries and its stability over time, the wage curve has been referred to as an 'empirical law of economics' (Card, 1995). There is substantial heterogeneity among wage curve analyses in terms of data sets, model specifications and econometric methods; the most prominent contributions include studies which look at Great Britain (Bell, Nickell and Quinitini, 2002; Johnes, 2007), Germany (Baltagi and Blien, 1998; Baltagi, Blien and Wolf, 2009), the United States (Blanchflower and Oswald, 2005), Australia (Kennedy and Borland, 2000) and New Zealand (Morrison, Papps and Poot, 2006). In the German context, we also see a growing literature which pays attention to, and corrects for, spatial dependence arising from cross-region interactions and spillovers (Buettner, 1999; Baltagi, Blien and Wolf, 2000; Elhorst, Blien and Wolf, 2007; Longhi, Nijkamp and Poot, 2006).

On the whole, the wage curve research prompted by Blanchflower and Oswald's work confirms the existence of a negative relationship between earnings and the rate of unemployment; this suggests imperfectly competitive labour markets where firms are not wage takers but adjust the level of pay downwards as local joblessness increases. However, the approach of investigators so far has been to attest the empirical validity of the wage curve simply by verifying whether the relationship is replicated in their data. This has resulted in a large amount of international evidence on the magnitude and significance of the wage curve elasticity, with almost one thousand estimates available at the time of Nijkamp and Poot's (2005) meta-analysis. Nevertheless, none of these authors has taken a step further to explore whether the wage curve can also be accepted as being superior to rival wage equations. To answer the question as to whether the wage curve truly is an 'empirical law of economics', the relationship needs being studied outside the confines of its own specific proposition and requires a directly confrontation with alternative earnings functions.

In this paper we go beyond model fitting and slope estimation to examine the relative success of the wage curve in the face of competing hypotheses of wage determination. In doing so, we look at Great Britain's 408 unitary authority and local authority areas (UALADs) over the period 1998-2010. The wage curve is thus tested against two contemporary theory-derived models which also provide an explanation for regional wage disparities, namely Urban Economics (UE) and New Economic Geography (NEG). Both have strong foundations in the urban and geographical economics literature, summarised by Huriot and Thisse (2000), Fujita and Thisse (2000), Fujita, Krugman and Venables (1999) and Brakman, Garretsen and Van Marrewijk (2009a), but propose distinct causes of pecuniary spatial externalities and economic agglomeration. Our study builds on Fingleton (2006, 2007), who evaluate the performance of NEG *vis-à-vis* UE using cross-sectional data respectively for Britain's UALADs and for the 200 NUTS2 regions of the European Union. We adopt the same approach to non-nested hypothesis testing; this involves estimating an Artificial Nested Model (ANM) which encompasses both models being compared (Davidson and MacKinnon, 1993; Hendry, 1995) as well as (a spatial version of) the J-test procedure (Kelejian, 2008; Burridge and Fingleton, 2010). We also allow for spatial effects that work through the disturbance term by specifying a spatially autoregressive error component structure, following the burgeoning spatial econometrics literature pioneered by Kapoor et al. (2008).

Of the two alternative rival theories, Urban Economics attributes a primary role to market linkages at the *intra-regional level*; productivity or wage variation is driven by differences in employment density, because of a greater variety of increasing-returns intermediate services accessible to final goods and services firms. By contrast, New Economic Geography emphasises market interdependencies at the *inter-regional level*; the wage rate increases with market potential because of cost advantages from locating close to large sources of supply and demand in high-income areas. As in UE, spatial concentration of economic activity reduces average production costs, allowing efficiency gains through greater exploitation of internal scale economies; in addition, proximity to final consumer markets and intermediate input providers implies lower transport costs, thus entailing pecuniary agglomeration externalities. Hence, Urban Economics and New Economic Geography have two distinct views on how economic geography can affect wages, respectively looking at within- and between-area market interactions, but there is no reference to labour market conditions (i.e. unemployment) in either theory.

The paper is organised as follows. Section 2 reviews the relevant wage curve literature and specify the model to be estimated empirically. Section 3 is concerned with the theoretical relationships coming from UE and NEG and the extended empirical specifications. Section 4 describes the variables and the data. Section 5 briefly sets out the FGS2SLS+GMM estimation strategy, while results are presented in Section 6 for the wage curve, UE and NEG models in isolation. In Section 7 the issue of comparing non-nested rival models is considered, and ANM and spatial J-test results discussed. Section 8 concludes.

2. THE WAGE CURVE MODEL

2.1. Review of the Wage Curve literature

The extensive literature which has developed in the last couple of decades on the wage curve postulates that, other things constant, employees who work in areas of high unemployment earn less than those working in areas of low unemployment. Blanchflower and Oswald (1990), using US and British micro-data, are among the first to derive an inverse relationship between the wage rate paid to individuals and the rate of unemployment in the local labour market. In their subsequent book (Blanchflower and Oswald, 1994a) they replicate this research with individual-level data for sixteen countries, and discover an unemployment elasticity of pay which is very similar across different nations and between different time periods, of approximately -0.10. This empirical regularity is also documented in Blanchflower and Oswald (1995), in which they provide international evidence drawn from their seminal book, and in Blanchflower and Oswald (2005), where they confirm the existence of the U.S. wage curve using modern American data. More specifically for the British case (see also Blanchflower and Oswald, 1994b), they examine data on approximately 175,000 workers from the General Household Surveys of 1973-1990, and estimate a slope of around -0.10 after controlling for regional fixed effects and the individual characteristics of workers; this finding is robust to the sample selected, the procedure used, the inclusion of a labour force participation variable, and to race, skill and gender.

In contrast to Blanchflower and Oswald (1994b), other investigators find that the elasticity of the wage curve varies across different categories. For instance, Card (1995), Baltagi and Blien (1998) and Baltagi et al. (2009) find that wages are more responsive to unemployment rate variation (hence the wage curve is more elastic) for men (see also Fingleton and Longhi, 2011) who tend to be employed in sectors with relatively higher entry and exit costs as opposed to women, and also for the young, the skilled, foreigners (Baltagi et al., 2009), non-union members and private sector workers (Card, 1995) all of whom tend to have relatively weak bargaining power. However there are some exceptions; for example, Baltagi et al. (2000) and Kennedy and Borland (2000) report that in Eastern Germany and Australia respectively the unemployment elasticity of female earnings is higher than that of male earnings. For the weak bargaining power groups, the link between unemployment and pay is evidently stronger because, in depressed labour markets, they have more difficulty than others finding alternative jobs when threatened by dismissal (e.g. in the event of an industrial dispute) or by lay-off (e.g. during a negative economic shock), therefore employers do not need to remunerate them so well.

A less elastic wage curve for union workers seems to contradict Blanchflower and Oswald's union-bargaining explanation that in a slack labour market a union would be more concerned about the number of unemployed members than higher wages for its employed members, which could lead to accepting a lower negotiated level of pay. Efficiency wage or labour turnover costs provide another non-competitive labour market explanation for why, at least in the short-run, wages tend to be lower in labour markets with higher unemployment. The argument is that, when unemployment is higher, firms face lower costs of replacing workers while the costs for employees of job losses or of voluntary quits are higher, therefore the wage that firms pay in excess of market-clearing to motivate workers or increase their productivity is lower. By contrast, at times of low unemployment there are more alternatives open to workers, and as a consequence employers must offer higher wages to retain workers or avoid shirking[1].

[1] Another theory is the labour-contract model, where firms and their workers are assumed to maximise joint utility, which implies that higher wages may be associated with higher contractual employment and therefore with lower unemployment rates. Sato (2000) proposes a search model showing that equilibrium wages and unemployment are driven by productivity differentials across local labour markets, with higher productivity being related to higher wages and lower unemployment.

One of the main contributions in the U.K. context is that of Bell et al. (2002). Their analysis is based on wages over the period 1976-1997 which are compositionally corrected in order to eliminate intra-region correlation and grouped data bias[2]. The correction is implemented in two stages; first they regress individual-level wages on individual (time-varying) characteristics, individual fixed effects, and region-specific time dummies. The dummy variables parameter estimates from the first-stage regression, which vary by regions and are constant across individuals within each region, are then treated as compositionally wage levels, having been adjusted for individual variables. These compositionally corrected wages enters the second-stage OLS regression as the dependent variable, and are explained by region-level variables (including the unemployment rate) and region-specific time trends. Using this approach, Bell et al. (2002) obtain a (short-run) elasticity of pay with respect to unemployment of -0.034. Later, Johnes (2007) focuses on a two-level model where the first level refers to time periods while the second level consists of a cross-classification between individuals and regions. With this set-up, he is able to simultaneously allow for time fixed effects and region plus individual random effects in order to accommodate grouped data effects as well as unobserved heterogeneity. Using 1992-2003 data from the British Household Panel Survey and instrumenting the local unemployment rate by its one-year lag, Johnes (2007) estimates an unemployment elasticity of -0.05. Hence British evidence suggests that while a wage curve appears to exist for Britain its magnitude does not accord with the 'empirical law' of -0.10. A similar conclusion is reached by Nijkamp and Poot (2005) in their meta-analysis of international evidence. They show that the wage curve is a robust empirical phenomenon but its corrected elasticity, after controlling for publication bias and the effect of unemployment on hours worked, is no more than -0.05. Moreover, they confirm that the role of unemployment in the determination of earnings is reduced when human capital variables such as education and job experience are also considered.

2.2. The empirical Wage Curve specification

The wage curve literature reviewed in Section 2.1 neglects the geography of interaction among regional economies; wages in one area are described as a function of local employment conditions alone, while the impact of proximate labour markets and thus the pay or joblessness rates nearby is ignored. There are wage curve studies, however, which do give attention to the implications of spatial spillover effects for the validity and strength of Blanchflower and Oswald's 'law', and we draw from these in order to motivate the empirical specification for the wage curve equation in this paper.

---

[2] The grouped nature of the sample data is a possible cause of cross-section positive dependence in the error term and downward biased standard errors. When data is collected at two hierarchical levels of aggregation, namely individual (wages and personal or job characteristics) and region (unemployment), a grouped data bias is introduced (Moulton, 1990) because individuals within the same region can be expected to share unmeasured as well as observable characteristics.

One argument in favour of a spatial approach to the wage curve estimation is the presence of spatial correlation in regional unemployment; the rate of unemployment in one region is commonly found to positively correlate with that in surrounding regions, with high -(low)- unemployment areas clustered in space as a result. This is a widely recognised phenomenon in labour market research (Manning, 1994; Molho, 1995), and there are several suggested reasons as to why it may occur. First, as workers can be expected to commute between home and workplace within functional local labour market areas, the formal administrative units often used for data collection do not reflect these functional areas but cut across them, thus a functional area may be represented in several formal areas. This means that variables such as the unemployment rate, which will tend towards equilibrium within a given functional area, may typically have similar values in proximate formal administrative areas. Concentrations of high or low unemployment may also be due to the spatial patterns of employment growth (labour demand), the spatial distribution of population characteristics such as job skills (labour supply), and the geography of house prices as disadvantaged workers seek cheaper accommodation. If unemployment rates in nearby districts directly determine the local wage rate but their effects are omitted, then the wage curve model will be misspecified and estimates of the unemployment elasticity of pay will be biased. An explicit way to take account of these effects is by adding the spatially lagged unemployment variable, as in Buettner (1999) and Longhi et al. (2006).

Longhi et al. (2006) estimate a wage curve for Western Germany by fixed-effects 2SLS using data for 327 regions over the period 1990-1997, and find a non-spatial, baseline parameter of -0.05 compared to almost -0.06 after accounting for spatial effects. The inclusion of spatial variables, while leaving results generally unchanged, enables the authors to uncover a range of spatial processes that should be accounted for in a proper wage curve specification. They show that Blanchflower and Oswald's relationship is stronger if regions are more isolated, because the mobility costs associated with a job change (commuting, migration, job search) are higher in less accessible, remote regions and thus the local labour supply is relatively inelastic. This means that employers can reduce wage rates at times of rising unemployment without fearing a potential move of workers to adjacent areas, a finding which can be interpreted as evidence that the wage curve arises from monopsonistic competition in local labour markets. Consistent with this hypothesis is the finding that the elasticity is larger (more negative) if employment opportunities in surrounding regions are tighter. Longhi et al. (2006) also demonstrate that wage rates are directly affected by accessibility and by the spatially lagged unemployment rate, with respectively positive and negative effects on local pay. In particular, the conclusion that employers have to pay higher wages in strongly interacting, agglomerated regions in order to retain workers or increase productivity (as opposed to dispersed, mostly rural regions) is in line with efficiency wage and labour turnover cost theories; it is also a feature of the wage curve analysis carried out by Morrison et al. (2006) in the New Zealand context, using the weighted average of inter-region road travel time as a measure of regional accessibility.

An additional source of model misspecification and incorrect (biased and inconsistent) estimates is the spatial correlation in wage rates. Buettner (1999) tests for spatial effects in the German wage curve by including the spatial lag of the dependent variable, and finds strong support for the hypothesis that wage rates in neighbouring districts exert an autonomous influence on local pay.

Another type of cross-section dependence involves the error term, being manifest as spatially autocorrelated residuals, typically reflecting one or more of shocks that are common to different areas, unobserved across-area spillovers and externalities stemming from inter-economy linkages, and omitted spatially autocorrelated explanatory variables. In this case the usual assumption of spherical disturbances underpinning conventional inferential methods is violated, therefore failure to account for positive spatial residual correlation leads to inflated t-ratios and unreliable inference. Existing wage curve studies correct for spatial correlation in the residuals by means of a common factor approach. For example, looking at Eastern Germany, Baltagi, Blien and Wolf (2000) use 2SLS estimation on variables in first differences as a way to wipe out (time-invariant) regional fixed effects, possibly related to a common history or locally available natural resources, which may cause dependencies among closely located areas. A similar framework is adopted by Elhorst, Blien and Wolf (2007), whose strategy is to also eliminate time-period fixed effects by taking the differences between the dependent and independent variables for each region and a reference region. Their method can be seen as equivalent to that in Pesaran (2006), which suggests estimating an augmented model with cross-section averages of the regressand and regressors acting as placeholders for unobserved common factors. Therefore, in their treatment of spatial residual correlation, they do not explicitly define the various connections between regions involved in the transmission of spatial effects via weights matrices and thus do not invoke a spatially lagged error as is typically employed in spatial econometrics. Likewise, in Buettner (1999) and Longhi et al. (2006), econometric specifications with a spatial lag in the error term are not considered.

In the present paper we follow the strand of the wage curve literature which deals with labour market interactions but we also rethink the error structure in order to explicitly model the spatial effects working through the disturbance term. Blanchflower and Oswald's relationship is thus extended by adding a spatially autoregressive error process (Kapoor et al., 2007), which assumes autocorrelated disturbances across space and time, to a random effects panel specification, together with spatial lags of the dependent and independent variables in order to capture the influence on local wages of the levels of pay or unemployment in areas nearby. The different spatial mechanisms that we introduce in the estimation of the wage curve can be summarised as follows:

$$
\begin{aligned}
&LogWage_t = \beta_0 + \rho(W * LogWage_t) + \beta_1 LogU_t + \theta(W * LogU_t) + \beta_2 LogT_t + \\
&+\beta_3 LogS_t + \beta_4 LogA_t + \omega_1 (LogU_t)(W * LogWage_t) + \omega_2 (LogU_t)(LogA_t) + e_t \\
&\qquad e_t = \lambda M e_t + \xi_t \\
&\qquad \xi_t = \mu + \nu_t
\end{aligned}
\tag{1}
$$

where the N x1 vector e_t is the spatially dependent error term, and this is a function of ξ_t which combines a permanent, that is time-constant, component $\mu \sim iid(0,\sigma_\mu^2)$, and a transient error component $\nu_t \sim iid(0,\sigma_\nu^2)$. **W** and **M** are N x N spatial weights matrices, defining the structure and intensity of linkages and spillovers among areas; these are specified in Section 4.

The variable A refers to an index of accessibility/agglomeration, and is computed following Longhi et al. (2006) as the sum of total employment Emp in all other districts weighted by distance, i.e. $A_{it} = \sum_j kEmp_{jt} d_{ij}^{-1}$ [3] [4].

A set of additional explanatory variables, consisting of (log) local schooling S and (log) local technical knowledge T as described in Section 4, is also included on the right-hand side of the earnings equation to control for composition effects. This is in line with applied wage curve studies using micro data which often consider regional averages of individual-level attributes such as gender, educational attainment, employment status (blue collar/white collar, part-time/full-time), firm size, industry classification, and other personal and job characteristics[5]. If the bargaining-power explanation for the existence of a wage curve holds, or if the monopsonistic power of employers in areas with a qualified and skilled workforce is lower as one would anticipate, then we might expect local wages to increase with T (β_2>0) and to decrease with S (β_3<0).

3. THE RIVAL MODELS

3.1. The model motivated by Urban Economics (UE) theory

This Section sets out the UE model, following Rivera-Batiz (1988), Abdel-Rahman and Fujita (1990), Ciccone and Hall (1996) and Fingleton (2003) [6]. The aim is to estimate a reduced form linking wage/productivity levels with density of economic activity in the form of employment density, and allowing for a direct test of the existence of increasing returns to agglomeration. UE theory emphasises the varying regional supply of non-traded producer services; in this set up market potential (i.e. transport costs, transport cost mediated price index variations and income variations across areas), which is at the core of NEG theory, is not relevant. Therefore spatial interdependencies and the rela-

[3] k is a scaling parameter, here equal to 10^{-6}.
[4] Following Longhi et al. (2006) we also construct variable A in a gravity-model fashion, allowing the agglomeration index of region i to depend both on its own employment size and that of its neighbours as in $A_{it} = \sum_j k(Emp_{jt} Emp_{it})^{0.5} d_{ij}^{-1}$. This alternative specification gives a similar set of estimation results.
[5] This is the approach commonly adopted after Blanchflower and Oswald (1994a, pp. 168-170) in order to avoid grouped data bias i.e. downward bias in the standard error of the unemployment coefficient (Moulton, 1990) and thus cross-section dependence in the error term (inter-region correlation).
[6] There are alternative UE models, e.g. Combes et al. (2008) and Brakman et al. (2009b), which lead to the same reduced form.

tive geographical position of regions do not matter, whereas local production conditions are crucial to explaining why some areas have higher wages than others.

The model assumes that the economy is divided into a traded sector (*M*), consisting of final goods and services produced under perfect competition, and a non-traded intermediate market service sector (*I*), characterised by monopolistic competition and providing inputs to the output (*Q*) of competitive industry *M*. So, assuming a Cobb-Douglas production function for *M*, we have

$$Q = (M^{\beta} I^{1-\beta})^{\alpha} L^{1-\alpha} = (M^{\beta} I^{1-\beta})^{\alpha} \tag{2}$$

where land (*L*) is equal to one because production is per unit area, *M* is the level of labour efficiency units employed in making *M*'s goods and services directly, and *I* is the level of producer services based on a CES sub-production function under monopolistic competition *à la* Dixit and Stiglitz (1977) that is

$$I = \left[\sum_{z=0}^{x} i_z^{1/\mu} \right]^{\mu} \tag{3}$$

As shown in the Appendix, this leads to a relationship between the level of final goods and services output *Q* and the total effective labour (in both final goods and services and intermediate market services) per unit area (*N*), as in

$$Q = (M^{\beta} I^{1-\beta})^{\alpha} = \phi N^{\gamma} \tag{4}$$

with constants ϕ and elasticity γ where

$$\gamma = \alpha[1 + (1-\beta)(\mu - 1)] \tag{5}$$

N is given by the product between total employment level per unit area (*E*) and each area's level of labour efficiency (*H*), i.e. $N=E \cdot H$. Parameter μ (μ>1) refers to internal increasing returns; it reflects the degree of product differentiation in the *I* sector, hence the strength of market power available to *I* firms, and defines the constant elasticity of substitution among service varieties $\sigma=\mu/(\mu-1)$ and the constant price elasticity of demand (see Appendix). Whether or not there are external increasing returns (**γ>1**) depends on the amount of internal scale economies being sufficiently large (**μ>1**), on the non-traded, increasing-returns sector being sufficiently important to final production (which is indexed by the magnitude of **β<1**), and on diminishing returns due to congestions costs[7] (**1-α<1**) being small enough so as not to outweigh the other two factors.

As long as γ>1, equation $Q = \phi N^{\gamma}$ captures increasing returns to density of economic activity given by *E* and *N* that are a result of the variety of produc-

[7] Crowding more and more workers onto the same unit area has detrimental effects on final output.

er services, which increases with clustering. Therefore efficiency gains (i.e. cost advantages) from internal increasing returns to scale in the producer service sector translate into productivity gains in the competitive, constant-returns final goods and services sector, due to agglomeration externalities i.e. external, city-wide increasing returns to scale; this is because final goods and services producers have a preference for a greater availability of producer service varieties which characterises larger towns and cities. Denser areas thus tend to have higher productivity and wage levels because of stronger local linkages between competitive industry and intermediate market service suppliers.

For the determination of wage rates, we use the equilibrium allocation of labour inputs to final production *Q*. This entails calculating the derivative of $Q = [f(N)]^{\alpha} L^{1-\alpha}$ with respect to *N*, which is the marginal product of labour, as follows

$$\begin{aligned} \mathrm{d}Q/\mathrm{d}N &= L^{1-\alpha} \cdot \alpha \cdot f(N)^{\alpha} \cdot f(N)^{-1} \\ &= f(N)^{\alpha} L^{1-\alpha} \cdot \alpha / f(N) \\ &= \alpha Q / f(N) \end{aligned} \tag{6}$$

We can replace *f(N)* with *N* without losing meaning and, following standard competitive equilibrium theory, we set the remuneration of effective labour (i.e. the nominal wage rate *w*) equal to its marginal product

$$w = \alpha Q / N \tag{7}$$

This implies that the share of final output going to labour is the wage rate per labour efficiency unit (*w*) times the number of labour efficiency units (*N*) divided by final output (*Q*), which is equal to the coefficient α as in $wN/Q = \alpha$. Taking the natural logarithm of both sides of $w = \alpha Q / N$ gives

$$Log(w) = Log(Q) + Log(\alpha) - Log(N) \tag{8}$$

Substituting for $Q = \phi N^{\gamma}$ and for *N*=*H·E* gives the short-run wage equation

$$\begin{aligned} Log(w) &= Log(\phi) + \gamma Log(H \cdot E) + Log(\alpha) - Log(H \cdot E) \\ &= k_1 + (\gamma - 1)Log(E) + (\gamma - 1)Log(H) \end{aligned} \tag{9}$$

in which k_1 denotes a constant. The estimated parameter for increasing returns to agglomeration is (*γ*-1) not *γ*, so it is possible to directly test for the presence of increasing returns by simply looking at the sign and significance of *(γ-1)*. In the absence of increasing returns, *γ=1* and the employment density variable disappears from the equation. When *γ>1* an increase in employment density yields a more than proportionate increase both in nominal wage rates (through the short-run wage equation) and in final output (since $Q = \phi N^{\gamma}$).

3.2. The model motivated by New Economic Geography (NEG) theory

The concept of market potential dates back to Harris (1954) but has long remained unmodelled because conventional assumptions of perfect competition and constant returns to scale could not provide a theoretical justification for the observed economic agglomeration phenomenon and the associated productivity advantages. Krugman (1991) was the first to develop a structural model around Harris's (1954) initial formulation, using the theory of monopolistic competition, product variety and internal scale economies introduced by Dixit and Stiglitz (1977). NEG theory thus has much in common with UE theory, having the same utility and profit maximising microfoundations and being based on the same market structure assumptions which give rise to pecuniary externalities from concentrated production.

Under Krugman's (1991) general equilibrium model (i.e. the basic NEG specification[8]), the wage equation is one of a set of simultaneous non-linear equations determining the equilibrium distribution of economic activity. This short-run relationship predicts that the nominal wage rate that firms in region *i* can afford to pay increases with market potential (*MP*) of region *i*, i.e. the level of access of region *i*'s firms to local and neighbouring markets.

$$Wage_i = \left[\sum_r Y_r (G_r)^{\sigma-1} (T_{ir})^{1-\sigma} \right]^{\frac{1}{\sigma}} = MP_i^{\frac{1}{\sigma}}$$
$$LogWage_i = \frac{1}{\sigma} LogMP_i \qquad (10)$$

Region *i*'s market potential depends on the level of income (*Y*) locally and in neighbouring areas, on trade costs (*T*) which increase with distance from *i*, and on the price index (*G*). A smaller elasticity of substitution among varieties (σ>1) tends to create a flatter market potential surface, whereas a larger elasticity of substitution leads to peaks centred on cities as the effect of distance is magnified as one moves away from accessible cities. Thus a smaller elasticity of substitution reduces the downward impact of distance on market potential and hence wages. Also, other regions will be less able to substitute for region *i*'s goods and services because a smaller σ (which also equals the price elasticity of demand) reduces the downward impact of competition on wages by increasing consumers' preference for product varieties, thus diminishing price competition from firms in other regions.

Krugman's (1991) formalised version of market potential is derived from economic theory but requires a number of pragmatic decisions in order to be estimated empirically. One advantage of Harris's (1954) original formulation is that it has less rigorous data requirements and does not necessitate stringent assumptions in order to be operationalised; he defines each region's market

[8] Extensions of the basic NEG specification include models developed by Krugman and Venables (1995) and Venables (1996), which incorporate UE-style input-output (vertical) linkages.

potential as the distance-weighted sum of market size (here proxied by population) of surrounding regions

$$MP_i = \sum_r Y_r d_{ir}^{-1} \tag{11}$$

Therefore, following Mion and Naticchioni (2005) and Brakman et al. (2009b), we draw from NEG theory to motivate the use of market potential while using Harris's definition to capture the extent of agglomeration externalities. The rationale is that the scope of the present analysis is not to structurally fit a NEG specification but to obtain a measure of the magnitude of NEG-style spatial linkages, without seeking to estimate and interpret the coefficient on the market potential variable as a function of the parameters of the underlying model. The use of Harris' market potential is supported by the finding in Head and Mayer (2004) that it performs fairly well when compared with a more structural measure.

3.3. The empirical extended UE and NEG specifications

Next we make some assumptions regarding the determinants of the variation in labour efficiency level (*H*) among areas. First we assume that *H* is affected by differences between workers in their ability to make productive use of the available technology, which is taken as homogenous across areas. We thus express the natural logarithm of each area's level of efficiency as a linear function of the level of educational attainment of resident workers; to avoid having to choose which level of schooling to consider, we focus on the percentage of working-age resident population with no qualifications. We denote this variable by ***S***. Another indicator of local area efficiency is the size of the local knowledge base. This variable is denoted by ***T***, and is approximated by the relative concentration of local employment in technology-intensive computing and R&D sectors (see Section 4 for a more precise description).

We also recognise that workers are mobile and wages paid at a workplace depend on the labour efficiency level at other locations from which workers commute; this means that the quality of the workforce at location *i* is determined by labour efficiency within commuting distance of *i*, as well as locally. Such efficiency spillovers are modelled via the term *W*LogH* which represents the matrix product of the standardised spatial weight matrix *W* (see Section 4) and the natural logarithm of *H*; more precisely, the contribution to region *i*'s efficiency level from in-commuting is given by row *i* of vector *W*LogH* which contains the sum of the weighted efficiency levels in other designated areas. This term captures the totality of the effects influencing local area efficiency, not only the effects of *S* and *T* locally but also their effects in other areas together with the local and remote effects of unmodelled factors represented by the random shocks[9].

[9] This is the outcome of a Leontief expansion, as shown in Fingleton (2006).

Combining the local (exogenous) variables assumed to influence an area's level of efficiency, we have

$$\begin{aligned} &Log(H) = b_0 + b_1 S + b_2 T + \rho W * Log(H) + \vartheta \\ &\vartheta \sim iid(0, \sigma_\vartheta^2) \end{aligned} \tag{12}$$

To find *W*LogH* in terms of known variables, we then rearrange $Log(w) = k_1 + (\gamma - 1)Log(E) + (\gamma - 1)Log(H)$ and multiply both sides by *W*. This gives

$$W * Log(H) = W * \frac{-k_1}{\gamma - 1} + \frac{1}{\gamma - 1} W * Log(w) - W * Log(E) \tag{13}$$

Substituting the expressions for *Log(H)* and *W*Log(H)* into $Log(w) = k_1 + (\gamma - 1)Log(E) + (\gamma - 1)Log(H)$, and adding an error term (ζ) which has the same spatial structure as that of the empirical wage curve model, we obtain:

$$\begin{aligned} Log(w) = k_1 + (\gamma - 1)Log(E) + (\gamma - 1)\Big[b_0 + b_1 S + b_2 T + \\ + \rho \left(W * \frac{-k_1}{\gamma - 1} + \frac{1}{\gamma - 1} W * Log(w) - W * Log(E) \right) + \vartheta \Big] + \varsigma \end{aligned} \tag{14}$$

Hence, simplifying and adding subscript *t*

$$\begin{aligned} &Log(w_t) = \beta_0 + \rho W * Log(w_t) + (\gamma - 1)(Log(E_t) - \rho W * Log(E_t)) + \\ &\quad + \beta_2 S_t + \beta_3 T_t + e_t \\ &\quad e_t = \lambda M e_t + \xi_t \\ &\quad \xi_t = \mu + \nu_t \end{aligned} \tag{15}$$

Using the same arguments, the empirical extended version of the NEG model is

$$\begin{aligned} &Log(w_t) = \beta_0 + \rho W * Log(w_t) + b_1 (Log(MP_t) - \rho W * Log(MP_t)) + \\ &\quad + \beta_2 S_t + \beta_3 T_t + e_t \\ &\quad e_t = \lambda M e_t + \xi_t \\ &\quad \xi_t = \mu + \nu_t \end{aligned} \tag{16}$$

4. VARIABLES AND DATA

The present Section provides a comprehensive description of the key variables which are considered in this study, all expressed in levels (before taking logs). These have been thoroughly presented or theoretically derived in previous chapters, where the Wage Curve, Urban Economics and New Economic Geography models are outlined.

The following table contains details on data sources, variable definitions and summary statistics for the variables in question, excluding transformations of these such as interaction terms and spatial lags. (Log) Wage spans the period 1998-2008 while all predictors, except (log) *S* which refers to year 2001, denote *N* x 1 vectors (for N=408 regions) at time *t* (with t=1999…2009). Regressors are thus lagged by one year so that they pre-date the period of analysis and can be treated as exogenous.

Data sources and variable definitions

	Source[a]	Description	Mean	Min	Max
Wage	Annual Survey of Hours and Earnings (ASHE)	**Mean Gross Weekly Wage Rate** (pay, in £ p/w, at the place of employment; all occupations, all persons)	£468.85	£166.64	£1,210.80
T	Annual Business Survey (ABS)	**Technical Knowledge** **S**eries of location quotients, i.e. measure of relative employment specialisation, in high knowledge-based sectors, namely computing & related activities and R&D, with LQ>1 high; LQ=1 none; LQ<1 low (local employment share in 1992 SICs 72 & 73 over the national share)	0.3	0.0	3.5
S	ONS 2001 Census	(Lack of) **Educational Attainment** (percentage of residents with no qualifications)	28.6%	10.0%	45.6%
U	ONS / JobCentre Plus	**Claimant Counts Ratio** (proportion of working-age population claiming unemployment-related benefits[b])	2.4%	0.2%	10.5%
E	Annual Business Survey (ABS)	**Employment Density** (total employment per square kilometre)	867	3	96,125[c]
MP	ONS mid-year population estimates	**Market Potential** (distance-weighted sum of population in adjacent areas, as in Eq. 11 of Ch. 3.2)	2,261	358	5,468

[a] *All data is available from NOMIS, the ONS' labour market statistics database.*
[b] *Since 1996 only people claiming Jobseeker's Allowance have been counted.*
[c] *City of London.*

The N x N standardised spatial weights matrix $\mathbf{W}$ takes the following form:

$$\begin{aligned}
&W_{ij}^{*} = \exp(-\hat{\sigma}_i d_{ij}) \text{ for } i \neq j \\
&W_{ij}^{*} = 0 \qquad\qquad\quad \text{ for } i = j \\
&W_{ij}^{*} = 0 \qquad\qquad\quad \text{ for } \mathrm{d}_{ij} > 100\text{km} \\
&W_{ij} = \frac{W_{ij}^{*}}{\sum_{j=1}^{N} W_{ij}^{*}}
\end{aligned} \qquad (17)$$

where $\hat{\sigma}_i$ is calibrated on commuting flows (Fingleton, 2003). It should be noted that the commuting data used to obtain $\mathbf{W}$ is taken from the UK's Census for the year 1991, therefore spatial weights are pre-determined with respect to wage data. The choice of a $\mathbf{W}$ matrix which pre-dates the dependent variable rules out potential concerns about the exogeneity of $\mathbf{W}$ and the consistency of estimates, by ensuring that causation can only run from commuting to pay. This $\mathbf{W}$ matrix is adopted to construct spatially lagged variables $W * LogWage$ and $W * LogU$, the latter referring to a Wage Curve specification. $W * LogWage$ is the spatial lag of local earnings, given by the product between $\mathbf{W}$ and the (log) wage vector at each time period t; the size and significance of the coefficient on this term indicates the responsiveness of a region's level of pay to variations in the wage rate of surrounding areas. $W * LogU$ is the spatial lag of the (log) unemployment rate vector; its inclusion allows testing the extent to which local wage variability can be explained by the unemployment rate in neighbouring regions.

The spatial weights matrix $\mathbf{M}$ for the error process is given by:

$$\begin{aligned}
&M_{ij}^{*} = 1 \text{ if } i \text{ and } j \text{ are contiguous i.e. share a border} \\
&M_{ij}^{*} = 0 \text{ otherwise} \\
&M_{ij} = \frac{M_{ij}^{*}}{\sum_{j=1}^{N} M_{ij}^{*}}
\end{aligned} \qquad (18)$$

5. ESTIMATION STRATEGY

This Section briefly sets out the FGS2SLS plus GMM procedure for estimating a random effects panel model with an endogenous spatial lag and spatially autoregressive disturbances (SARAR-RE Model).

There are three estimation stages. Stages 1 and 3 provide estimates of $b = (\beta_0, \rho, \beta_1, \theta, \beta_2, \beta_3, \beta_4, \omega_1, \omega_2)$ for the wage curve model,

$b = (\beta_0, \rho, (\gamma - 1), \beta_2, \beta_3)$ for UE or $b = (\beta_0, \rho, b_1, \beta_2, \beta_3)$ for NEG, but differ in the values adopted for σ_ν^2, σ_1^2 and λ; Stage 1 uses arbitrary values of 1, 1 and 0 respectively for σ_ν^2, σ_1^2 and λ, but estimates are available from the data via Stage 2 for use in Stage 3.

In both Stages 1 and 3, error dependence is eliminated using a Cochrane-Orcutt (C-O) transformation to give ξ. This is done by pre-multiplying by $(I_{TN} - \lambda I_T \otimes M)$ since $e = (I_{TN} - \lambda I_T \otimes M)^{-1}\xi$ and $\xi = (I_T \otimes (I_N - \lambda M))e$, thus

$$\begin{aligned} Y^* &= (I_T \otimes (I_N - \lambda M))Y \\ X^* &= (I_T \otimes (I_N - \lambda M))X \end{aligned} \tag{19}$$

The TN x $f \geq (k$+1) matrix of instruments Z comprises a linearly independent subset of the exogenous variables, and matrices X and Z are assumed to be full column rank with $f \geq (k$+1). The error covariance matrix is $\Omega_\xi = E(\xi\xi') = \sigma_u^2(J_T \otimes I_N) + \sigma_\nu^2 I_{TN}$, which means that the disturbances are non-spherical, and therefore $P_Z = Z(Z'\hat{\Omega}_\xi Z)^{-1}Z'$, which is a symmetric matrix as $P_Z\Omega_\xi$ is idempotent. The vector of regression coefficients is therefore

$$\begin{aligned} \hat{b} &= \left[(X^{*\prime}Z)(Z'\hat{\Omega}_\xi Z)^{-1}(Z'X^*) \right]^{-1} (X^{*\prime}Z)(Z'\hat{\Omega}_\xi Z)^{-1}(Z'Y^*) = \\ &= (X^{*\prime}P_Z X^*)^{-1} X^{*\prime}P_Z Y^* \end{aligned} \tag{20}$$

and the estimated variance-covariance matrix is given by

$$\hat{C} = \left[(X^{*\prime}Z)(Z'\hat{\Omega}_\xi Z)^{-1}(Z'X^*) \right]^{-1} = (X^{*\prime}P_Z X^*)^{-1} \tag{21}$$

Greene (2003) refers to equivalent equations as generalized methods of moments (instrumental variables) estimators with non-spherical disturbances.

The standard errors of the $\hat{b}$ are given by the squares roots of the values on the main diagonal of $\hat{C}$, which allows 't-ratios' to be calculated for purposes of inference.

Stage 2 consists of GMM estimation, based on Kapoor et al. (2007). First the relationships between unknown population vectors and matrices is defined

$$\Gamma\phi' - \eta = 0 \text{ and } \tilde{\Gamma}\phi' - \tilde{\eta} = 0 \tag{22}$$

where Γ and $\tilde{\Gamma}$ are 3 x 4 matrices, η and $\tilde{\eta}$ are 3 x 1 vectors and the vector of parameters to be estimated is $\phi = \begin{bmatrix} \lambda & \lambda^2 & \sigma_v^2 & \sigma_1^2 \end{bmatrix}$. With the estimated disturbances $\hat{e}$ given by Stage 1, the estimates of unknown vectors η and $\tilde{\eta}$, denoted by g and $\tilde{g}$ can be obtained, that is

$$\begin{aligned} g &= \hat{e}' Q_0 \hat{e} \\ \tilde{g} &= \hat{e}' Q_1 \hat{e} \end{aligned} \tag{23}$$

The corresponding sample counterparts of matrices Γ and $\tilde{\Gamma}$ given by G and $\tilde{G}$ can be obtained also. These sample counterparts are defined as follows

$$G = \begin{bmatrix} \frac{2}{N(T-1)} \hat{e}' Q_0 \hat{e}_{-1} & \frac{-1}{N(T-1)} \hat{e}'_{-1} Q_0 \hat{e}_{-1} & 1 & 0 \\ \frac{2}{N(T-1)} \hat{e}'_{-2} Q_0 \hat{e}_{-1} & \frac{-1}{N(T-1)} \hat{e}'_{-2} Q_0 \hat{e}_{-2} & \frac{1}{N} t_1 & 0 \\ \frac{1}{N(T-1)} (\hat{e}' Q_0 \hat{e}_{-2} + \hat{e}'_{-1} Q_0 \hat{e}_{-1}) & \frac{-1}{N(T-1)} \hat{e}'_{-1} Q_0 \hat{e}_{-2} & 0 & 0 \end{bmatrix} \tag{24}$$

$$G \begin{bmatrix} \lambda & \lambda^2 & \sigma_v^2 & \sigma_1^2 \end{bmatrix}' - g = \zeta(\lambda \quad \sigma_v^2 \quad \sigma_1^2) \tag{25}$$

$$\tilde{G} = \begin{bmatrix} \frac{2}{N} \hat{e}' Q_1 \hat{e}_{-1} & \frac{-1}{N} \hat{e}'_{-1} Q_1 \hat{e}_{-1} & 0 & 1 \\ \frac{2}{N} \hat{e}'_{-2} Q_1 \hat{e}_{-1} & \frac{-1}{N} \hat{e}'_{-2} Q_1 \hat{e}_{-2} & 0 & \frac{1}{N} t_1 \\ \frac{1}{N} (\hat{e}' Q_1 \hat{e}_{-2} + \hat{e}'_{-1} Q_1 \hat{e}_{-1}) & \frac{-1}{N} \hat{e}'_{-1} Q_1 \hat{e}_{-2} & 0 & 0 \end{bmatrix} \tag{26}$$

$$\tilde{g} = \begin{bmatrix} \frac{1}{N} \hat{e}' Q_1 \hat{e} \\ \frac{1}{N} \hat{e}'_{-1} Q_1 \hat{e}_{-1} \\ \frac{1}{N} \hat{e}' Q_1 \hat{e}_{-1} \end{bmatrix} \tag{27}$$

$$\tilde{G}\begin{bmatrix}\lambda & \lambda^2 & \sigma_\nu^2 & \sigma_1^2\end{bmatrix}' - \tilde{g} = \tilde{\zeta}(\lambda \quad \sigma_\nu^2 \quad \sigma_1^2) \tag{28}$$

in which $t_1 = tr(M'M)$, $\hat{e}_{-1} = (I_T \otimes M)\hat{e}$, $\hat{e}_{-2} = (I_T \otimes M)\hat{e}_{-1}$, and $\xi(\lambda \quad \sigma_\nu^2 \quad \sigma_1^2)$ and $\hat{\xi}(\lambda \quad \sigma_\nu^2 \quad \sigma_1^2)$ are residuals. The parameter estimates are given by

$$\begin{aligned}(\hat{\lambda}, \hat{\sigma}_\nu^2, \hat{\sigma}_1^2) = \arg\min \big\{ &\zeta(\lambda \quad \sigma_\nu^2 \quad \sigma_1^2)'\zeta(\lambda \quad \sigma_\nu^2 \quad \sigma_1^2) + \\ &+ \tilde{\zeta}(\lambda \quad \sigma_\nu^2 \quad \sigma_1^2)'\tilde{\zeta}(\lambda \quad \sigma_\nu^2 \quad \sigma_1^2) \big\}\end{aligned} \tag{29}$$

These can be obtained via non-linear least squares estimation. Given that the variance of the two components within argmin{…} are not the same, Kapoor et al. (2007) suggest differential weighting. However for simplicity equal weight is given to each of the six moments equations, which also gives consistent estimates.

6. ESTIMATION RESULTS

6.1. The spatial Wage Curve model

6.1.1. Parameter estimates

Table 1.a reports the results of fitting wage equations motivated by the wage curve literature. The (log) unemployment rate and (log) agglomeration are lagged by one year and thus predetermined, whereas the spatial lag of *LogWage* is clearly endogenous because of multilateral spatial dependence between the wage observations[10]. Therefore we instrument *W*LogWage* by the first-order spatial lag of covariates *LogA*, *LogT* and *LogS* following Kelejian and Prucha (1998); the other instruments are the exogenous variables *LogU*, *LogA*, *LogT* and *LogS*. The starting point is a specification as eq. 1 in Section 2.2. This model (Table 1.a, first column) gives an unemployment elasticity of pay that has the expected negative sign but a p-value (0.2104) which exceeds conventional Type I error rates. Its spatial lag, however, is highly significant and closely resembles the 'law' of -0.10. The other parameter estimates are all appropriately signed and statistically significant except for the interaction terms. In particular, the coefficient on the interaction term involving *LogU* and *W*LogU* is negative, thus correctly suggesting that the wage curve is more elastic in areas surrounded by high unemployment, but has no additional explanatory power. With regards to the interaction of *LogU* and *LogA*, not only is this statistically irrelevant but it is also wrongly signed in the light of the expectation that more accessible areas should have a less elastic wage curve, as was suggested by Longhi et al. (2006).

[10] Wages observed in any two regions *i* and *j* are correlated with the error term and, since there is two-way causation between them through the spatial lag of *LogWage*, *W*LogWage* also correlates with the error term.

Table 1.a. Results from the spatial Wage Curve model estimated in isolation (*dependent variable*: *LogWage*)

	1	2	3	4
	FGS2SLS +GMM	**FGS2SLS +GMM (ρ = 0)**	**FGS2SLS +GMM ($\beta_4 = \omega_1 = \omega_2 = 0$)**	**FGS2SLS+ GMM ($\theta = \beta_4 = \omega_1 = \omega_2 = 0$)**
Spatial Externalities/ Monopsony Effects				
$W * LogWage_{it}$ (ρ)	**0.0793**		**0.0876**	**0.0427**
(t-stat)	(4.41)***		(5.51)***	(2.38)***
Local Unemployment				
$LogU_{it}$ (β_1)	**-0.1552**	**-0.4393**	**0.0185**	**-0.0203**
(t-stat)	(-1.13)	(-3.83)***	(1.47)*	(-1.67)**
Unemployment within Commuting Distance				
$W * LogU_{it}$ (θ)	**-0.1260**	**-0.1118**	**-0.1480**	
(t-stat)	(-4.49)***	(-3.93)***	(-6.67)***	
Interaction of $LogU_{it}$ with its *Spatial Lag*				
$(LogU_{it})$ X $(W * LogU_{it})$ (δ_1)	**-0.0106**	**-0.0486**		
(t-stat)	(-0.48)	(-2.30)***		
Agglomeration/Accessibility				
$LogA_{it} = Log(\sum_j TotEmp_{jt} / Dist_{ij})$ (β_4)	**0.1276**	**0.1787**		
(t-stat)	(4.45)***	(7.15)***		
Interaction of $LogU_{it}$ with $LogA_{it}$				
$(LogU_{it})$ X $(LogA_{it})$ (δ_2)	**-0.0243**	**-0.0698**		
(t-stat)	(-1.23)	(-4.35)***		
Local knowledge				
$LogT_{it}$ (β_2)	**0.0369**	**0.0415**	**0.0419**	**0.0559**
(t-stat)	(8.59)***	(9.37)***	(9.17)***	(11.50)***
Local Unskilled Workforce				
$LogS_i$ (β_3)	**-0.1070**	**-0.1098**	**-0.1539**	**-0.1172**
(t-stat)	(-3.59)***	(-3.38)**	(-4.71)***	(-3.59)***
Constant	7.2165	8.0865	6.4751	6.6494
(t-stat)	(26.88)***	(43.32)***	(41.85)***	(40.27)***
Error process				
λ [a]	**0.6628*** **	**0.6511*** **	**0.6686*** **	**0.6278*** **
σ_v^2	0.0035	0.0038	0.0037	0.0047
$\sigma_1^2 = \sigma_v^2 + T\sigma_u^2$	0.0561	0.0604	0.0613	0.0685
RSS	75.70	83.68	83.94	99.02
R^{2} [b]	0.7389	0.7035	0.7169	0.6546
No. areas	408	408	408	408
No. in-sample years (1999-2009)	11	11	11	11

[a] *Standard error (not reported) of the spatial autoregressive parameter is obtained by bootstrapping.*
[b] *Correlation between observed and fitted values of LogWage.*

We see strong evidence of a spatially autoregressive process involving both *LogWage* and the errors. As suggested by $\hat{\rho}$ = 0.0793 with a t-ratio of 4.41, the spatially weighted wages in surrounding areas located within commuting distance have a positive and significant effect on local wages. First, the endogenous spatial lag may be picking up the benefits associated with proximity to high-productivity, high-wage areas (external scale economies from urban density). Secondly, higher earnings in nearby areas may be raising the opportunity wage of local workers, thereby increasing the wage rate that local employers need to pay in order to attract or retain workers. Hence, while *W*LogWage* in the urban economics and economic geography models comes from an auxiliary SAR process involving labour efficiency (Section 3.3), here it represents something different i.e. spatial externalities and/or monopsonistic competition in local labour markets. We also find significantly positive spatial correlation in the disturbance term ($\hat{\lambda}$ = 0.6628), which points to the presence of global shock effects being transmitted across the urban hierarchy with feedback loops.

In the specification without an endogenous spatial lag (second column), the coefficient on *LogU* becomes significant but its size does not seem to be appropriate, being well below the empirical regularity of -0.10. Given our results so far, it is apparent that the spatial lag of the (log) wage rate should be in the model while there is not much evidence of significant interaction variables. Moreover, it can be argued that *LogA* should be omitted from the wage curve specification because it may be measuring the same proximity effects as Market Potential, and we want to maintain a clear distinction between the wage curve model and the rival (non-nested) NEG specification. This leads us to the third specification, which reaffirms the main finding from model 1, namely that local wages evidently do not respond to unemployment in the immediate area but are strongly determined by unemployment within commuting distance. Importantly, the presence of *W*LogWage* in the earnings equation requires a special interpretation of the estimated coefficients, one which is different from ordinary regression. We undertake this task below in order to give a precise account of the impact of unemployment, agglomeration and the locational variables on wages. It should be noted that, in a specification without the spatial lag of *LogU* (column 4), local unemployment becomes negatively signed although it does not reach the two-tailed level of significance; accordingly, our preferred specification remains model 3 (after excluding *LogU*). Moreover model 4 is evidently omitting important spatial effects, as reflected by a lower correlation between actual and predicted wages.

6.1.2. Marginal effects

LeSage and Pace (2009) point out that, when the endogenous spatial lag is in the model, the true total effect on a dependent variable Y of a unit change in an exogenous variable X_k is not the same as the regression coefficient estimate β_k , because the true partial derivative $\partial Y / \partial X_k$ also takes account of changes passing through the simultaneous dependence system.

The regression in column 3 of Table 1.a is equivalent to a *spatial Durbin model* (Elhorst, 2010a, 2010b) with spatially autoregressive error components (time-constant region-specific random effects plus time- and region-varying disturbances). In this specification, regional variation in (log) wage levels depends also on the level of pay in neighbouring regions, as captured by *W*LogWage*, and on the rate of unemployment in neighbouring regions, as represented by *W*LogU*. The model thus accommodates multi-regional interdependencies up and down the spatial network, and expands the information set for the *i*th region to include observations on the dependent and explanatory variables in other regions. The general implication of including the spatial lags of the regressand and regressor is that a change in *LogU* associated with a given region *i* will directly affect *LogWage* in region *i* itself but will potentially have an indirect impact on *LogWage* in all other regions also. This is different from non-spatial linear regression (based on the assumption of independence among cross-sectional units) where $\partial y_i / \partial x_{ik} = \beta_k$ for all *i* while $\partial y_i / \partial x_{jk} = 0$ for $i \neq j$.

The proper interpretation of the marginal effects of the *k*th explanatory variable can be derived from the panel data specification

$$Y_t = \rho W Y_t + \alpha + X_t \beta + W X_t \theta + e_t$$

$$(I - \rho W) Y_t = \alpha + X_t \beta + W X_t \theta + e_t \tag{30.a}$$

$$Y_t = (I - \rho W)^{-1} \alpha + (I - \rho W)^{-1} (X_t \beta + W X_t \theta) + (I - \rho W)^{-1} e_t$$

$$(I - \rho W)^{-1} = I + \rho W + \rho^2 W^2 + \rho^3 W^3 + \ldots \tag{30.b}$$

from which it follows that, at a given time *t*, the matrix of partial derivatives of *Y* in the different regions (y_i for *i*=1,…, *N*) with respect to X_k in the different regions (x_{ik} for *i*=1,…, *N*) varies over *i* and is equal to

$$S_k = \frac{\partial Y}{\partial X_k} = \begin{bmatrix} \dfrac{\partial y_1}{\partial x_{1k}} & \cdots & \dfrac{\partial y_1}{\partial x_{Nk}} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial y_N}{\partial x_{1k}} & \cdots & \dfrac{\partial y_N}{\partial x_{Nk}} \end{bmatrix} = (I - \rho W)^{-1} \begin{bmatrix} \beta_k & w_{12}\theta_k & \cdots & w_{1N}\theta_k \\ w_{21}\theta_k & \beta_k & \cdots & w_{2N}\theta_k \\ \vdots & \vdots & \beta_k & \vdots \\ w_{N1}\theta_k & w_{N2}\theta_k & \cdots & \beta_k \end{bmatrix} \tag{31}$$

or $S_k = (I - \rho W)^{-1} C_k$. These differentiated effects can be summarised by summing the total effects over the rows (or columns) of the matrix S_k and then computing the mean over all regions, as in

$$N^{-1} \sum_{ij}^{N} \frac{\partial Y_i}{\partial X_{jk}} = N^{-1} \iota' \left[(I - \rho W)^{-1} C_k \right] \iota \tag{32}$$

It is possible to distinguish the *average total effects* of a unit change in X_k on *Y* between two types of impact. The *average row effect* quantifies the *average total impact* <u>*to*</u> *an observation*; this is the mean of the elements of an N x 1 column vector, where each element is the sum of the impacts on a particular unit of observation (region) of the dependent variable of a unit change in *all* x_{ik}

i.e. the kth explanatory variable across all N regions. The *average column effect* quantifies the *average total impact* *from an observation*; this is the mean of the elements of a 1 x N row vector, where each element is the sum of the impacts on *all* y_i resulting from a unit change in a particular unit of observation (region) of the kth explanatory variable.

This average total effect can be partitioned into a direct and indirect component. The *average direct effect* is a scalar summary measure of the own-partial derivatives, where each of these derivatives is the own-region (direct) impact i.e. the impact of a unit change in region i's X_k on region i's Y. It is calculated as the average of the main diagonal elements of the asymmetric N x N matrix S_k, as in

$$N^{-1}\sum_{j}^{N}\frac{\partial Y_j}{\partial X_{jk}} = N^{-1}trace\left[(I-\rho W)^{-1}I\beta_k\right] \qquad (33)$$

The *average indirect effect* is a scalar summary that corresponds to the cross-partial derivatives, or other-region (indirect) impacts, associated with a unit change in the explanatory variable. It thus represents the response of region i's Y to a unit change in X_k in all other regions. It is equal to the difference between the average total effect and the average direct effect, and is computed as the average of either the row sums or the column sums of the off-diagonal elements of matrix S_k.

Table 1.b. Direct, Indirect and Total Effects of Wage Curve (Model 3)'s variables

	$LogU_{it}$	$LogT_{it}$	$LogS_i$
AVG DIRECT (OWN-REGION) EFFECTS	**0.0176**	**0.0420**	**-0.1540**
(bootstrapped t-ratio)	(1.26)	(10.27)***	(-4.57)***
AVG INDIRECT (SPATIAL) EFFECTS	**-0.1591**	**0.0040**	**-0.0146**
(bootstrapped t-ratio)	(-7.15)***	(59.19)***	(-24.64)***
AVERAGE TOTAL EFFECTS	**-0.1415**	**0.0459**	**-0.1686**
(bootstrapped t-ratio)	(-5.36)***	(11.25)***	(-4.99)***

Table 1.b displays the true marginal effects related to the explanatory variables of the log wage rate (column 3), i.e. the average total impacts, and their average direct and indirect components. Interestingly, for *LogT* and *LogS*, it is the case that the spatial impacts are highly statistically significant but quantitatively small, so that the total impacts are mostly due to the own-region effects. By contrast, for *LogU*, local unemployment evidently has no influence on local wages, since the average own-region effects are not significant (as well as wrongly signed), but what counts for local wages is unemployment within commuting distance; therefore, the true wage curve elasticity amounts to -0.14, which is close to the 'empirical law' of -0.10, and this is effectively entirely attributable to other-region effects (and most likely monopsonistic local labour

markets). It should be noted that the direct effect is different from β_1, and the indirect effect associated with *LogU* is different from θ, because they incorporate some feedback loop effects, as implied by eq. 30.b; these arise because any given region is considered a neighbour to its neighbour, so impacts pass through neighbouring areas and eventually come back to the area of origin itself.

6.2. The Urban Economics (UE) model

Table 2 summarises the outcome of estimating the rival urban economics model by FGS2SLS, using the exogenous spatial lags *W*LogT* and *W*LogS* as instruments for the spatially lagged dependent variable (column 3). The (log) employment density predates the (log) wage rate by one year, thus it can be taken as exogenous. The estimates given in column 1 correspond to a basic UE specification, without controlling for local and in-commuting labour efficiency. The coefficient of 0.0348 on *LogE* (here ρ is constrained to the value 0) is highly significant and, quantitatively, means that doubling the number of employees per square kilometre increases local wages by 2.4% ($2^{0.0348}$-1=0.0244); this result is somewhat below the range of elasticities (3-8%) typically found in the agglomeration literature (Rosenthal and Strange, 2004). Allowing for efficiency variations (column 2) reduces the size of the coefficient by around one third, an indication that the labour efficiency variables are not orthogonal to employment density, although evidence remains of a very significant employment density effect. Commuting, as embodied in the endogenous spatial lag introduced in column 3, also emerges as a significant determinant of local wages ($\hat{\rho}$ = 0.0890 with t-ratio=5.49), causing the level of labour efficiency and thus the wage rate to be higher in workplace areas with intense in-commuting flows of qualified and skilled workers (in line with the interpretation of the endogenous lag as derived in Section 3.3).

Despite these significant labour efficiency spillovers, the efficiency level of the resident workforce is still a significant explanatory factor, with both variables *LogT* and *LogS* measured within each local area being statistically relevant and carrying the expected sign. In particular, doubling a region's relative specialisation in computing and R&D activities (*T*) raises wages by 2.24% ($2^{0.0320}$-1=0.0224), while a fall by a half in the proportion of working-age population without qualifications (*S*) produces an increase in wages by 13.45% ($2^{0.1821}$-1=0.1345).

A separate source of higher wages is represented by increasing returns to employment density; the estimated value of $(\gamma - 1)$ is positive and statistically significantly above zero, thus suggesting external scale economies due to pecuniary externalities boosting productivity and wages in areas with greater concentration of economic activity. More specifically, the preferred model (column 3) gives a density elasticity of 2.15% ; this is larger than other related (cross-sectional) studies of spatial wage disparities in Great Britain, e.g. Fingleton (2003, 2006), who find an urban wage premium on density of around 1.5%. That for France is even higher, at 3-4% (Barde, 2009), also after controlling for the effects of spatial sorting by individual skills (Combes et al., 2008), as captured by worker fixed effects, age and squared age.

Table 2. Results from the competing UE model estimated in isolation (*dependent variable*: *LogWage*)

	1	2	3
	FGS2SLS+GMM	**FGS2SLS+GMM**	**Iterated** [b] **FGS2SLS+GMM**
In-commuting Labour Efficiency $W*LogWage_{it}$ (ρ) (t-stat)			**0.0890** (5.49)***
Employment Density $(LogE_{it} - \rho W*LogE_{it})$ $(\gamma-1)$ (t-stat)	**0.0348** (7.87)***	**0.0199** (5.28)***	**0.0215** (5.96)***
Local Knowledge Base $LogT_{it}$ (β_2) (t-stat)		**0.0507** (10.79)***	**0.0320** (6.73)***
Local Unskilled Workforce $LogS_i$ (β_3) (t-stat)		**-0.1774** (-6.94)***	**-0.1821** (-7.68)***
Constant (t-stat)	5.9408 (208.78)***	6.9562 (82.88)***	6.3030 (47.16)***
Error process			
λ [a]	**0.6898*** **	**0.5354*** **	**0.5397*** **
σ_v^2	0.0039	0.0059	0.0052
$\sigma_1^2 = \sigma_v^2 + T\sigma_u^2$	0.0973	0.0692	0.0608
RSS	138.12	101.27	88.97
$R^{2\,c}$	0.4320	0.6270	0.7031
No. areas	408	408	408
No. in-sample years (1999-2009)	11	11	11

[a] *Standard error (not reported) of the spatial autoregressive parameter is obtained by bootstrapping.*
[b] *Iteration is to satisfy the constraint involving ρ.*
[c] *Correlation between observed and fitted values of LogWage.*

Importantly, as pointed out in Corrado and Fingleton (2012), the partial derivative $\partial LogWage/\partial LogE$ is simply equal to $(\gamma-1)$ i.e. the coefficient estimate on $(LogE - \rho WLogE)$, or $(I-\rho W)LogE$, so the presence of the spatial lag *W*LogWage* does not require any special interpretation of the marginal effect of the compound employment density variable. This is because the empirical specification of the urban economics model, which includes the endogenous spatial lag, is a reduced form resulting from an auxiliary SAR process involving (log) labour efficiency (Section 3.3).

The positive sign and statistical significance of the estimated λ reflects positive spatial residual correlation, hence positive dependence among the permanent and transient error components, hinting at spatially autocorrelated heterogeneity and omitted variables. Moreover, with a correlation between actual and predicted values of log wages equal to 70%, the UE model (column 3) has a level of the fit which is almost the same as that of the wage curve model, which is an informal indication that neither model is quantitatively superior to the other. We carry out a more formal analysis subsequently.

Table 3. Results from the competing NEG model estimated in isolation (*dependent variable*: *LogWage*)

	1	2	3
	FGS2SLS+GMM	**FGS2SLS+GMM**	**Iterated [b] FGS2SLS+GMM**
In-Commuting Labour Efficiency			
$W * LogWage_{it}$ (ρ)			**0.0635**
(t-stat)			(4.11)***
Market Potential			
$(LogMP_{it} - \rho W * LogMP_{it})$ (ϕ)	**0.2137**	**0.1225**	**0.1236**
(t-stat)	(6.85)***	(6.04)***	(5.93)***
Local Knowledge Base			
$LogT_{it}$ (β_2)		**0.0507**	**0.0411**
(t-stat)		(12.32)***	(10.14)***
Local Unskilled Workforce			
$LogS_i$ (β_3)		**-0.1343**	**-0.1278**
(t-stat)		(-5.53)***	(-5.54)***
Constant	4.4933	5.9806	5.5578
(t-stat)	(18.76)***	(31.93)***	(27.72)***
Error process			
λ [a]	**0.7185*** **	**0.5825*** **	**0.5867*** **
σ_v^2	0.0036	0.0052	0.0049
$\sigma_1^2 = \sigma_v^2 + T\sigma_u^2$	0.0830	0.0617	0.0580
RSS	124.56	97.55	89.17
R^2 [c]	0.4980	0.6415	0.6856
No. Areas	408	408	408
No. in-sample years (1999-2009)	11	11	11

[a] *Standard error (not reported) of the spatial autoregressive parameter is obtained by bootstrapping.*
[b] *Iteration is to satisfy the constraint involving ρ.*
[c] *Correlation between observed and fitted values of LogWage.*

6.3. The New Economic Geography (NEG) model

Table 3 gives results of estimating the empirical NEG model. The coefficient on the composite market potential variable (column 1) is highly significant and positively signed, which is a strong indication of pecuniary external economies from geographical proximity to large nearby markets for intermediate inputs and final goods (as implied by lower transport costs, and lower average production costs). The market potential effect is noticeably smaller once the technology and schooling variables are controlled for (column 2), thus suggesting its correlation with *LogT* and *LogS*.

It should be noted that our indicators of local area efficiency possibly pick up non-pecuniary externalities, which operate through non-market interactions and depend on the technological and skill content of local employment. The significance and appropriate sign of the coefficients on *LogT* and *LogS* might thus indicate the presence of 'technological externalities' involving localised knowledge spillovers[11]. These considerations also hold for the UE results discussed previously.

There is evidence that in-commuting labour efficiency effects are also important (column 3), as inferred from a spatially autoregressive coefficient which is statistically significant and positively signed ($\hat{\rho}$ =0.0635, t-ratio=4.11). In this specification, the composite market potential variable retains its quantitative size and explanatory power, which means that neighbouring effects driven by commuting are specifically identified and captured by the endogenous spatial lag, separately from market potential effects.

According to results in column 3, a 1% increase in market potential is associated with a wage improvement of 0.12%. This estimate is smaller than those derived by other spatial panel studies also applying a FGS2SLS+GMM procedure but structurally estimating a short-run NEG wage equation. Fingleton (2008) uses a panel of 77 countries in the years 1970, 1980, 1990 and 2000 to fit a model with spatially and temporally autocorrelated disturbances, also controlling for educational attainment measures (i.e. years of schooling and the literacy ratio) and a time trend, and obtains an elasticity of 0.45. With the same empirical specification (except for the time trend) but at a lower spatial scale of analysis, Amaral et al. (2010) tests the relationship between market potential and nominal wages for Brazilian municipalities over the period 1980-2000, and arrives at a coefficient of 0.35; his approach is the same as that adopted by Fingleton (2008) also in the choice of instruments, namely the exogenous schooling and literacy variables as well as the absolute latitude of each geographical unit and its square. Within the GB context, the study of Great Britain's local authority areas carried out by Fingleton (2006) gives a value of 0.15, which is very close to that found here[12]; his paper mainly differs from the present analysis in that it is based on cross-sectional observations and does not incorporate spatially autoregressive errors, but the empirical assumptions regarding local labour efficiency variation are similar.

[11] Technological externalities usually refer to external economies from access to a large pool of skilled workers and learning externalities from information flows. Cross-sector knowledge spillovers (external to the firm and industry but internal to the city) stem from industrial diversification and are referred to as urbanisation or Jacobian externalities (Jacobs, 1969), whereas own-sector knowledge spillovers (external to the firm but internal to the industry) are referred to as localisation or Marshallian externalities and may be due to the higher degree of beneficial specialisation (Marshall, 1890) or of innovative activity (Arrow, 1962) possible in denser areas.

[12] Evidence of positive pecuniary externalities stemming from proximity to large markets has been found by Hanson (1998) in the United States and Mion (2004) in Italy (see also Roos, 2001 and Brakman *et al.*, 2004 for Germany, and Niebuhr, 2006 for 158 European regions).

7. MODELS COMPARISON AND SELECTION

One important issue is that the wage curve and the competing UE or NEG theory are non-nested models, because the explanatory variables of one are not a subset of the explanatory variables of the other, therefore constraining the relevant parameters to zero does not reduce from one to the other. This means that it is not possible to simply restrict parameters and use such tests as the Likelihood Ratio in order to decide between non-nested rival hypotheses.

7.1. Results obtained using 'Inclusive Regressions'

To shed light on which model might be the preferred specification from an econometric perspective, we initially adopt an 'inclusive regression' approach (Davidson and MacKinnon, 1993; Hendry, 1995). Thus, after estimating the models individually, we combine the wage curve and either the UE or the NEG theory in a single empirical specification of which each non-nested rival hypothesis is a special case. Fingleton (2006, 2007) refer to this composite data generating process (DGP) as an artificial nested model (ANM). We then look at whether, even in the presence of the competing variable, (log) unemployment retains its significance and the elasticity that one would anticipate under the wage curve. Since the DGP nests the unemployment hypothesis and thus explains the data generated by the wage curve, if either UE or NEG encompasses the DGP then we can infer that Blanchflower and Oswald's relationship is encompassed by the competing model.

We find that neither the employment density hypothesis (Table 4) nor the market potential hypothesis (Table 5) dominates the wage curve; that is, unemployment does not lose its predictive power when directly confronted by each of these separate effects. The coefficients on the UE variable in Table 4 and on the NEG variable in Table 5 are also significantly above zero, an indication that reducing from the ANM to the wage curve by restricting either of these parameters to zero is not feasible as the ANM's fit would be significantly lowered; this is evidence that, given the presence of *LogU* or *W*LogU* (together with *LogS*, *LogT* and *W*LogWage*), intra- and inter-region economic geography play an additional role in predicting wages. At this stage we are not able to say which one, if any, of the rival UE and NEG hypotheses is more challenging for the wage curve; however the consensus among the empirical geographical economics literature is that, at a low level of spatial aggregation, market potential has weaker explanatory force while employment density is more relevant (Brakman et al., 2009b; Brülhart and Mathys, 2008), therefore we may expect UE to be the stronger competing paradigm.

So far we have seen that both unemployment and either employment density or market potential should enter the earnings equation. This is the case irrespective of whether local unemployment or unemployment within commuting distance is considered, although our results confirm that specifications with the latter variable should be preferred, since the estimated coefficient on *W*LogU* is more in line with the 'empirical law' of -0.10 (consistently with Table 1.a).

Table 4. Results from 'inclusive regressions' nesting Wage Curve and UE models (*dependent variable*: *LogWage*)

	1	2	3	4
	FGS2SLS+ GMM (ρ = 0)	**FGS2SLS+ GMM (ρ = 0)**	**Iterated FGS2SLS+ GMM**	**Iterated FGS2SLS+ GMM**
Endogenous Spatial Lag				
$W * LogWage_{it}\ (\rho)$			**0.1083**	**0.1185**
(t-stat)			(6.53)***	(7.58)***
Local Unemployment				
$LogU_{it}\ (\beta_1)$	**-0.0648**		**-0.0722**	
(t-stat)	(-4.84)***		(-5.70)***	
Unemployment within Commuting Distance				
$W * LogU_{it}\ (\theta)$		**-0.1272**		**-0.1262**
(t-stat)		(-6.24)***		(-6.52)***
Employment Density				
$(LogE_{it} - \rho W * LogE_{it})\ (\gamma - 1)$	**0.0292**	**0.0178**	**0.0326**	**0.0197**
(t-stat)	(6.56)***	(4.37)***	(7.72)***	(5.23)***
Local Knowledge Base				
$LogT_{it}\ (\beta_2)$	**0.0510**	**0.0489**	**0.0309**	**0.0268**
(t-stat)	(10.35)***	(9.86)***	(6.46)***	(5.63)***
Local Unskilled Workforce				
$LogS_i\ (\beta_3)$	**-0.0797**	**-0.1362**	**-0.0769**	**-0.1538**
(t-stat)	(-2.39)***	(-4.96)***	(-2.55)***	(-6.18)***
Constant	6.6246	6.9013	5.8223	6.0877
(t-stat)	(60.60)***	(76.37)***	(36.93)***	(46.25)***
Error process				
λ [a]	**0.5783*** **	**0.5913*** **	**0.6063*** **	**0.6165*** **
σ_v^2	0.0047	0.0043	0.0040	0.0036
$\sigma_1^2 = \sigma_v^2 + T\sigma_u^2$	0.0743	0.0751	0.0611	0.0618
RSS	95.35	90.95	80.56	76.52
R^2 [b]	0.6563	0.6759	0.7382	0.7498
No. Areas	408	408	408	408
No. in-sample years (1999-2009)	11	11	11	11

[a] *Standard error (not reported) of the spatial autoregressive parameter is obtained by bootstrapping.*
[b] *Correlation between observed and fitted values of LogWage.*

Moreover we see that (log) schooling and (log) knowledge are statistically significant and appropriately signed, a result which endorses an extended model as set out in Section 3.3 that incorporates local labour efficiency variations. The spatially lagged dependent variable also has significant explanatory power, and improves the level of fit noticeably when is added to the ANM.

However it leaves estimates broadly unchanged, and this result is somewhat different from existing evidence; for example, Fingleton (2006) finds that spatial (commuting) effects nullify the impact of market potential in the ANM specification. In a later study seeking to explain individual-level wages from the

British Household Panel Database, Fingleton and Longhi (2011) estimate an ANM which combines per-district, within-commuting-distance employment density and market potential and additionally controls for local unemployment as well as a set of individual-level covariates (e.g. age, marriage, children); they also find that, having taken spatial effects into account, market potential is not a factor affecting pay. Another important result is that the relative performance of the wage curve and UE models depends on gender; that is, in the presence of the alternative wage predictors, unemployment is significant for male respondents but not for females while the reverse is observed for the spatially weighted employment density variable.

Table 5. Results from 'inclusive regressions' nesting Wage Curve and NEG models (*dependent variable*: *LogWage*)*

	1	2	3	4
	FGS2SLS+ GMM (ρ = 0)	**FGS2SLS+ GMM (ρ = 0)**	**Iterated FGS2SLS+ GMM**	**Iterated FGS2SLS+ GMM**
Endogenous Spatial Lag				
$W*LogWage_{it}$ (ρ)			**0.0688**	**0.1026**
(t-stat)			(4.38)***	(7.06)***
Local Unemployment				
$LogU_{it}$ (β_1)	**-0.0219**		**-0.0253**	
(t-stat)	(-1.89)**		(-2.25)***	
Unemployment within Commuting Distance				
$W*LogU_{it}$ (θ)		**-0.1223**		**-0.1247**
(t-stat)		(-5.86)***		(-6.09)***
Market Potential				
$(LogMP_{it} - \rho W*LogMP_{it})$ (ϕ)	**0.1229**	**0.1186**	**0.1228**	**0.1154**
(t-stat)	(5.58)***	(4.94)***	(5.41)***	(4.51)***
Local Knowledge Base				
$LogT_{it}$ (β_2)	**0.0532**	**0.0485**	**0.0437**	**0.0360**
(t-stat)	(11.92)***	(11.38)***	(10.15)***	(9.10)***
Local Unskilled Workforce				
$LogS_i$ (β_3)	**-0.0951**	**-0.0985**	**-0.0831**	**-0.1059**
(t-stat)	(-2.98)***	(-3.87)***	(-2.76)***	(-4.52)***
Constant	5.8776	5.9571	5.4204	5.3868
(t-stat)	(28.36)***	(27.70)***	(24.64)***	(23.79)***
Error process				
λ [a]	**0.6115*** **	**0.6556*** **	**0.6196*** **	**0.6808*** **
σ_v^2	0.0047	0.0038	0.0044	0.0035
$\sigma_1^2 = \sigma_v^2 + T\sigma_u^2$	0.0648	0.0638	0.0601	0.0552
RSS	96.24	87.78	87.26	75.80
R^2 [b]	0.6479	0.6861	0.6950	0.7410
No. Areas	408	408	408	408
No. in-sample years (1999-2009)	11	11	11	11

[a] *Standard error (not reported) of the spatial autoregressive parameter is obtained by bootstrapping.*
[b] *Correlation between observed and fitted values of LogWage.*

Table 6.a. Direct, Indirect and Total Effects of Table 4 variables (inclusive model nesting Wage Curve and UE)

	$LogU_{it}$	$LogT_{it}$	$LogS_i$
AVG DIRECT (OWN-REGION) EFFECTS	**-0.0369**	**0.0275**	**-0.1089**
(bootstrapped t-ratio)	(-2.41)***	(5.84)***	(-3.46)***
AVG INDIRECT (SPATIAL) EFFECTS	**-0.1140**	**0.0038**	**-0.0149**
(bootstrapped t-ratio)	(-5.53)***	(50.43)***	(-33.99)***
AVERAGE TOTAL EFFECTS	**-0.1509**	**0.0313**	**-0.1238**
(bootstrapped t-ratio)	(-5.45)***	(6.66)***	(-3.96)***

Table 6.b. Direct, Indirect and Total Effects of Table 5 variables (inclusive model nesting Wage Curve and NEG)

	$LogU_{it}$	$LogT_{it}$	$LogS_i$
AVG DIRECT (OWN-REGION) EFFECTS	**0.0088**	**0.0348**	**-0.1223**
(bootstrapped t-ratio)	(0.87)	(8.85)***	(-4.59)***
AVG INDIRECT (SPATIAL) EFFECTS	**-0.1451**	**0.0039**	**-0.0138**
(bootstrapped t-ratio)	(-7.22)***	(69.47)***	(-36.70)***
AVERAGE TOTAL EFFECTS	**-0.1362**	**0.0388**	**-0.1362**
(bootstrapped t-ratio)	(-5.54)***	(9.83)***	(-5.10)***

Our next task, given the presence of the endogenous spatial lag in model 4 of Tables 4 and 5, is to also consider the direct, indirect and total effects of our variables, as previously done in Section 6.1.2 for the wage curve model. Results from estimating 'inclusive regressions' nesting the wage curve and the UE model (Table 6.a) and the wage curve and the NEG model (Table 6.b) are obtained including both *LogU* and *W*LogU* in the fitted specification, comparably to Table 1.b in Section 6.1.2. With regards to $(LogE - \rho W * LogE)$ and $(LogMP - \rho W * LogMP)$, as mentioned earlier these compound variables do not require any special interpretation of the marginal effect, and so we focus on the direct and indirect effects of the main variables of our basic wage curve equation, *LogU*, *LogT* and *LogS*.

Tables 6.a and 6.b differ from Table 1.b in terms of the inclusion of either $(LogE - \rho W * LogE)$ or $(LogMP - \rho W * LogMP)$ in the ANM specification, a presence which leaves estimated effects broadly unchanged. It turns out that, when

Blanchflower and Oswald's law is challenged by the UE theory, the average direct impact of *LogU* becomes negative and significant, as one would expect under the Wage Curve, a finding which may be related to the estimated coefficient on employment density being economically less important (with an elasticity of around 0.02) than that on market potential (with elasticity of about 0.l2). Overall though, when one considers the average total effect of *LogU*, it is not dissimilar under either model, and also of a similar order of magnitude to the classic Blanchflower and Oswald's law.

7.2. Results obtained using the spatial J-test

The evidence we have presented so far suggests that the goodness of fit of the wage curve model is about the same as that produced by models derived from the Urban Economics or New Economic Geography theories, and that the unemployment hypothesis is not dominated by, nor dominates, the employment density or market potential hypotheses. We next provide further evidence about the relative performance of the wage curve, UE theory and NEG theory using the J-test procedure (Davidson and MacKinnon, 1981, 1982).

This allows us to test a null model, Model_0, against an alternative, and non-nested, model, Model_1. Here the testing problem involves first estimating the H1: UE model or H1: NEG model to obtain fitted values of $LogWage_{UE}$ or fitted values of $LogWage_{NEG}$, and then adding these as an auxiliary variable to the maintained H0: Wage Curve model. We then test whether the predictive value from H1 has significant explanatory power given the presence of the wage curve variables. To accomplish this, we adopt a spatial extension due to Kelejian (2008) of Davidson and MacKinnon's (1981) J-test, which allows specifications containing spatial lags in both the dependent variable and the disturbance term, namely the SARAR-RE model.

7.2.1. Test specification

Under the null hypothesis, SARAR-RE Model_0 is true

$$\begin{aligned} \mathbf{Y} &= \rho_0 \mathbf{WY} + \mathbf{X}_0 \boldsymbol{\beta}_0 + \mathbf{e}_0 \\ \mathbf{e}_0 &= \lambda_0 \mathbf{Me}_0 + \boldsymbol{\xi}_0 \end{aligned} \tag{34}$$

where $\mathbf{Y}$ is the N x 1 vector of observations on *LogWage*, $\mathbf{X}_0$ is the N x k_0 matrix of observations on exogenous regressors *LogU* (or *WLogU*), *LogT* and *LogS*, $\mathbf{W}$ and $\mathbf{M}$ are the non-stochastic pre-defined matrices of exogenous spatial weights, $\mathbf{e}_0$ is the N x 1 vector of disturbance terms, and $\xi_0 \sim (\boldsymbol{\iota}_T \otimes \mathbf{I}_N)\boldsymbol{u}_0 + \boldsymbol{v}_0$ is the unobserved shock vector with time constant region effects picking up regional heterogeneity, $\boldsymbol{u}_0 \sim iid(0, \sigma_{u0}^2)$, and innovations $\boldsymbol{v}_0 \sim iid(0, \sigma_{v0}^2)$. The parameters to be estimated are the slope coefficients in the k_0 x 1 vector $\boldsymbol{\beta}_0$, the spatial autoregressive parameters ρ_0 and λ_0, and the error variances σ_{v0}^2 and σ_{u0}^2. Under the alternative, the data are generated by a similar structure, giving SARAR-RE Model_1

$$\begin{aligned}\mathbf{Y} &= \rho_1 \mathbf{WY} + \mathbf{X}_1 \boldsymbol{\beta}_1 + \mathbf{e}_1 \\ \mathbf{e}_1 &= \lambda_1 \mathbf{Me}_1 + \boldsymbol{\xi}_1\end{aligned} \tag{35}$$

in which $\mathbf{X}_1$ is the N x k_1 matrix of observations on the competing explanatory variable *LogE* (under the rival UE hypothesis) or *LogMP* (under the rival NEG hypothesis) plus the other exogenous variables *LogT* and *LogS*. The spatial autoregressive processes involving $\mathbf{Y}$ and $\mathbf{e}$, and the exogenous weighting matrices $\mathbf{W}$ and $\mathbf{M}$ which govern them, are identical to those in Model_0. Kelejian (2008) extends his approach to allow for a finite number $g \geq 1$ of non-nested alternatives, but in the interest of simplicity and clarity, we opt to keep the UE and NEG rivals as two separate competing hypotheses and consider each of these in turn.

As in Burridge and Fingleton (2010), we write $\mathbf{X}_0 = [\mathbf{X}_{00} \vdots \mathbf{X}_{01} \vdots \mathbf{X}_{02} \vdots \mathbf{X}_{03}]$ and $\mathbf{X}_1 = [\mathbf{X}_{10} \vdots \mathbf{X}_{11} \vdots \mathbf{X}_{12} \vdots \mathbf{X}_{13}]$, in which $\mathbf{X}_{00} = \mathbf{X}_{10} = [1,1,...,1]'$ and the remaining *X*s are the exogenous regressors in the null and alternative models. Also, $\mathbf{Z}_0 = [\mathbf{WY} \vdots \mathbf{X}_{00} \vdots \mathbf{X}_{01} \vdots \mathbf{X}_{02} \vdots \mathbf{X}_{03}]$, $\boldsymbol{\alpha}_0 = [\rho_0, \boldsymbol{\beta}_0']'$, $\mathbf{Z}_1 = [\mathbf{WY} \vdots \mathbf{X}_{10} \vdots \mathbf{X}_{11} \vdots \mathbf{X}_{12} \vdots \mathbf{X}_{13}]$ and $\boldsymbol{\alpha}_1 = [\rho_1, \boldsymbol{\beta}_1']'$, so that the following notation can be adopted for the null and alternative models

$$\begin{array}{lll} H_0 : \textit{WAGE CURVE} & \mathbf{Y} = \rho_0 \mathbf{WY} + \mathbf{X}_0 \boldsymbol{\beta}_0 + \mathbf{e}_0 & = \mathbf{Z}_0 \alpha_0 + \mathbf{e}_0 \\ H_1 : \textit{UE} & \mathbf{Y} = \rho_{UE} \mathbf{WY} + \mathbf{X}_{UE} \boldsymbol{\beta}_{UE} + \mathbf{e}_{UE} & = \mathbf{Z}_{UE} \alpha_{UE} + \mathbf{e}_{UE} \\ H_1 : \textit{NEG} & \mathbf{Y} = \rho_{NEG} \mathbf{WY} + \mathbf{X}_{NEG} \boldsymbol{\beta}_{NEG} + \mathbf{e}_{NEG} & = \mathbf{Z}_{NEG} \alpha_{NEG} + \mathbf{e}_{NEG} \end{array} \tag{36}$$

We implement the test in four steps. *Step 1* consists of estimating α_0 and α_1 in eq. 36 by instrumental variables (IV) since, given the presence of $\mathbf{WY}$, the ordinary least squares (OLS) estimator would be inconsistent. As in Kelejian (2008) we define the following matrices, respectively for Model_0, Model_1 (which can be either UE or NEG) and the model that combines both (for later use). We thus have

$$\mathbf{L}_0 = [\mathbf{X}_0 \vdots \mathbf{WX}_{02} \vdots \mathbf{WX}_{03}] = [\mathbf{X}_{00} \vdots \mathbf{X}_{01} \vdots \mathbf{X}_{02} \vdots \mathbf{X}_{03} \vdots \mathbf{WX}_{02} \vdots \mathbf{WX}_{03}] \tag{37.a}$$

$$\mathbf{L}_1 = [\mathbf{X}_1 \vdots \mathbf{WX}_{12} \vdots \mathbf{WX}_{13}] = [\mathbf{X}_{10} \vdots \mathbf{X}_{11} \vdots \mathbf{X}_{12} \vdots \mathbf{X}_{13} \vdots \mathbf{WX}_{12} \vdots \mathbf{WX}_{13}] \tag{37.b}$$

$$\mathbf{L}_{01} = [[\mathbf{X}_0 \vdots \mathbf{X}_1] \vdots \mathbf{WX}_{02} \vdots \mathbf{WX}_{03} \vdots \mathbf{WX}_{12} \vdots \mathbf{WX}_{13}] \tag{37.c}$$

It should be noted that, contrary to Kelejian (2008), we are excluding $\mathbf{WX}_{01}$ and $\mathbf{WX}_{11}$ because our main variables of interest (namely unemployment within commuting distance, and employment density and market potential when ρ is not constrained to zero) are spatially lagged by definition (i.e. $W*LogU$) or already include a spatial lag as they are empirically derived as being net of spatial (commuting) effects (i.e. $(LogE-\rho W*LogE)$ and $(LogMP-\rho W*LogMP)$). Secondly, we use a minimal set of instruments, as advocated by Burridge and Fingleton (2010), by only adopting first-order $\mathbf{W}$ matrices, although Kelejian (2008) allows $\mathbf{W}$s of any order up to arbitrary small integer *r*. Thirdly, covariates are the same, thus $\mathbf{WX}_{02}=\mathbf{WX}_{12}=W*LogT$ and $\mathbf{WX}_{03}$ = $\mathbf{WX}_{13}$ = *W*LogS*. Next we construct the matrices of instruments $\mathbf{H}_0=\left[\mathbf{L}_0\vdots\mathbf{M}\,\mathbf{L}_0\right]_{LI}$, $\mathbf{H}_1=\left[\mathbf{L}_1\vdots\mathbf{M}\,\mathbf{L}_1\right]_{LI}$ and $\mathbf{H}_{01}=\left[\mathbf{L}_{01}\vdots\mathbf{M}\,\mathbf{L}_{01}\right]_{LI}$ in which the subscript *LI* denotes a spanning set of linearly independent columns. These instrument sets, from which we exclude $\mathbf{MX}_{01}$ and $\mathbf{MX}_{11}$ consistently with the arguments laid out above, give the projection matrices $\mathbf{P}_0=\mathbf{H}_0(\mathbf{H}_0'\mathbf{H}_0)^{-1}\mathbf{H}_0'$ and $\mathbf{P}_1=\mathbf{H}_1(\mathbf{H}_1'\mathbf{H}_1)^{-1}\mathbf{H}_1'$, leading to the IV estimators for the null and alternative models respectively

$$\hat{\alpha}_{0,\mathrm{IV}}=\left[\mathbf{Z}_0'\mathbf{P}_0\mathbf{Z}_0\right]^{-1}\mathbf{Z}_0'\mathbf{P}_0\mathbf{Y} \tag{38.a}$$

$$\hat{\alpha}_{1,\mathrm{IV}}=\left[\mathbf{Z}_1'\mathbf{P}_1\mathbf{Z}_1\right]^{-1}\mathbf{Z}_1'\mathbf{P}_1\mathbf{Y} \tag{38.b}$$

In *Step 2* the vectors of residuals from IV estimation of the null and alternative models, defined as $\hat{\mathbf{e}}_0=\mathbf{Y}-\mathbf{Z}_0\hat{\alpha}_0$ and $\hat{\mathbf{e}}_1=\mathbf{Y}-\mathbf{Z}_1\hat{\alpha}_1$, are used to estimate λ_0 and λ_1, σ_{0u}^2 and σ_{0v}^2, σ_{1u}^2 and σ_{1v}^2 via the non-linear GMM method of Kapoor et al. (2007) explained in Section 5.

In *Step 3*, we use $\hat{\lambda}_0$ and $\hat{\lambda}_1$ to construct the spatially lag-transformed variables

$$\begin{aligned}
\boldsymbol{Y}^*(\hat{\lambda}_0)&=(\boldsymbol{I}-\hat{\lambda}_0\mathbf{M})\boldsymbol{Y}\\
\mathbf{Z}_0^*(\hat{\lambda}_0)&=(\boldsymbol{I}-\hat{\lambda}_0\mathbf{M})\mathbf{Z}_0\\
\boldsymbol{Y}^*(\hat{\lambda}_1)&=(\boldsymbol{I}-\hat{\lambda}_1\mathbf{M})\boldsymbol{Y}\\
\mathbf{Z}_1^*(\hat{\lambda}_1)&=(\boldsymbol{I}-\hat{\lambda}_1\mathbf{M})\mathbf{Z}_1
\end{aligned} \tag{39}$$

Then, using the same instrument sets $\mathbf{H}_0$ and $\mathbf{H}_1$ as previously, and the estimated covariance matrices $\hat{\mathbf{\Omega}}_{0\zeta}=\hat{\sigma}_{0u}^2(\boldsymbol{J}_T\otimes\boldsymbol{I}_N)+\hat{\sigma}_{0v}^2\boldsymbol{I}_{TN}$ and for the disturbance terms $\hat{\mathbf{\Omega}}_{1\zeta}=\hat{\sigma}_{1\mu}^2(\boldsymbol{J}_T\otimes\boldsymbol{I}_N)+\hat{\sigma}_{1v}^2\boldsymbol{I}_{TN}$, we estimate by IV the resulting equations for, respectively, the null and alternative hypotheses

$$
\begin{aligned}
(\boldsymbol{I}-\hat{\lambda}_0\mathbf{M})\boldsymbol{Y} &= (\boldsymbol{I}-\hat{\lambda}_0\mathbf{M})(\mathbf{Z}_0\alpha_0+\mathbf{e}_0) \\
\mathbf{Y}^*(\hat{\lambda}_0) &= \mathbf{Z}_0^*(\hat{\lambda}_0)\alpha_0+\mathbf{e}_0^*(\hat{\lambda}_0)
\end{aligned}
\tag{40.a}
$$

$$
\begin{aligned}
(\boldsymbol{I}-\hat{\lambda}_1\mathbf{M})\boldsymbol{Y} &= (\boldsymbol{I}-\hat{\lambda}_1\mathbf{M})(\mathbf{Z}_1\alpha_1+\mathbf{e}_1) \\
\mathbf{Y}^*(\hat{\lambda}_1) &= \mathbf{Z}_1^*(\hat{\lambda}_1)\alpha_1+\mathbf{e}_1^*(\hat{\lambda}_1)
\end{aligned}
\tag{40.b}
$$

With $\mathbf{P}_0^* = \left[\mathbf{H}_0(\mathbf{H}_0'\hat{\mathbf{\Omega}}_{0\xi}\mathbf{H}_0)^{-1}\mathbf{H}_0'\right]$ and $\mathbf{P}_1^* = \left[\mathbf{H}_1(\mathbf{H}_1'\hat{\mathbf{\Omega}}_{1\xi}\mathbf{H}_1)^{-1}\mathbf{H}_1'\right]$, we thus obtain parameter estimates, fitted values and the residual vector for the Wage Curve model

$$\hat{\alpha}_{0,FGS2SLS} = \left[\mathbf{Z}_0^{*\prime}\mathbf{P}_0^*\mathbf{Z}_0^*\right]^{-1}\mathbf{Z}_0^{*\prime}\mathbf{P}_0^*\mathbf{Y}^*(\hat{\lambda}_0) \tag{41.a}$$

$$\hat{\mathbf{Y}}^*(\hat{\lambda}_0) = \mathbf{Z}_0^*(\hat{\lambda}_0)\hat{\alpha}_{0,FGS2SLS} \tag{41.b}$$

$$\hat{\mathbf{e}}_0(\hat{\lambda}_0) = \mathbf{Y}-\hat{\mathbf{Y}}^*(\hat{\lambda}_0) \tag{41.c}$$

and the rival, either UE or NEG, model

$$\hat{\alpha}_{1,FGS2SLS} = \left[\mathbf{Z}_1^{*\prime}\mathbf{P}_1^*\mathbf{Z}_1^*\right]^{-1}\mathbf{Z}_1^{*\prime}\mathbf{P}_1^*\mathbf{Y}^*(\hat{\lambda}_1) \tag{42.a}$$

$$\hat{\mathbf{Y}}^*(\hat{\lambda}_1) = \mathbf{Z}_1^*(\hat{\lambda}_1)\hat{\alpha}_{1,FGS2SLS} \tag{42.b}$$

$$\hat{\mathbf{e}}_1(\hat{\lambda}_1) = \mathbf{Y}-\hat{\mathbf{Y}}^*(\hat{\lambda}_1) \tag{42.c}$$

In *Step 4*, we augment the right-hand side of the Wage Curve in eq. 39.a with predictions $\hat{\mathbf{Y}}^*(\hat{\lambda}_1)$ from eq. 41.a, which approximate the forecast value of the competing (UE or NEG) theory, to obtain

$$
\begin{aligned}
\mathbf{Y}^*(\hat{\lambda}_0) &= \mathbf{Z}_0^*(\hat{\lambda}_0)\alpha_0+\hat{\mathbf{Y}}^*(\hat{\lambda}_1)\delta+\mathbf{e}_0^*(\hat{\lambda}_0) \\
&= \mathbf{Z}_{01}^{**}(\hat{\lambda}_0)\eta+\mathbf{e}_{01}^{**}(\hat{\lambda}_0)
\end{aligned}
\tag{43}
$$

Under the null model, $\eta_0 = [\alpha_0', 0]'$ is assumed to be the true value but we estimate $\eta = [\alpha_0', \delta]'$; that is, we test the Wage Curve model (the null model) against either the rival UE theory or the rival NEG theory (the alternative model) in terms of the hypotheses H_0: $\delta=0$ against H_1: $\delta\neq 0$. Following Kelejian (2008), we use the instrument matrix $\mathbf{H}_{01}$ as defined previously, which is equal to $\mathbf{H}_{01} = \left[\left[\mathbf{X}_0 \vdots \mathbf{X}_1\right] \vdots \mathbf{WX}_{02} \vdots \mathbf{WX}_{03} \vdots \mathbf{MX}_{02} \vdots \mathbf{MX}_{03} \vdots \mathbf{MWX}_{02} \vdots \mathbf{MWX}_{03}\right]_{LI}$ since $\mathbf{X}_{02} = \mathbf{X}_{12}$ and $\mathbf{X}_{03} = \mathbf{X}_{13}$. This gives the projection matrix $\mathbf{P}_{01}^{**} = \left[\mathbf{H}_{01}(\mathbf{H}_{01}'\hat{\mathbf{\Omega}}_{0\xi}\mathbf{H}_{01})^{-1}\mathbf{H}_{01}'\right]$ so that the FGS2SLS estimator of η is

$$\hat{\eta}_{FGS2SLS} = \left[\mathbf{Z}_{01}^{**\prime} \mathbf{P}_{01}^{**} \mathbf{Z}_{01}^{**} \right]^{-1} \mathbf{Z}_{01}^{**\prime} \mathbf{P}_{01}^{**} \mathbf{Y}^{*} (\hat{\lambda}_0) \tag{44}$$

and the estimated variance-covariance matrix of the slope coefficients is

$$\hat{\mathbf{V}} = \left[(\mathbf{Z}_{01}^{**\prime} \mathbf{H}_{01})(\mathbf{H}_{01}{}' \hat{\mathbf{\Omega}}_{0\xi} \mathbf{H}_{01})^{-1} (\mathbf{H}_{01}{}' \mathbf{Z}_{01}^{**}) \right]^{-1} = (\mathbf{Z}_{01}^{**\prime} \mathbf{P}_{01}^{**} \mathbf{Z}_{01}^{**})^{-1} \tag{45}$$

which is used to construct a Wald test statistic for δ=0 in eq. 43. More precisely, when the null is true

$$\hat{J} = \frac{(\hat{\eta}(l))^2}{\hat{\mathbf{V}}(l,l)} \rightarrow^{d} \chi_1^2 \tag{46}$$

where l is the number of elements in η, so that $\hat{\eta}(l)$ is the last parameter i.e. δ, and $\hat{\mathbf{V}}(l,l)$ is its estimated variance.

7.2.2. Test results

Tables 7.a and 7.b present the outcome of the spatial J-test procedure for discriminating between non-nested models. In Table 7.a, we treat the wage curve as the null or maintained model, testing the local unemployment hypothesis first (columns 1 and 2) and its spatial lag then (columns 3 and 4) against either of the employment density or market potential hypotheses. The null that the fitted values from the UE or NEG models add no explanatory information given the wage curve is rejected in all cases, as the J-statistic is significantly larger than the critical value from the reference χ_1^2 distribution under the null. We can therefore reject the wage curve as a complete explanation *per se* because of the additional explanatory power of the non-nested rival theories. We have also evidence of relatively weaker rejection of the wage curve when H_0: *W*LogU* is considered as the maintained hypothesis; this is expected since, as we showed previously, unemployment within commuting distance has stronger statistical and economic significance than local unemployment.

Having seemingly rejected the wage curve, this does not imply that UE and NEG hold. This becomes evident when we treat each of these in turn as the maintained hypothesis. According to results in Table 7.b, neither UE nor NEG are acceptable *per se*, since we find that the wage curve is capable of falsifying both of these competing models, given that we always reject the null that the coefficient on the fitted values from the wage curve is zero (i.e. the null that the wage curve has no additional predictive power given UE or NEG). Interestingly, we also see that the J-statistic leads to relatively stronger rejection when H_1: *W*LogU* is used as the alternative hypothesis. This adds support to the thesis that unemployment within commuting distance rather than local joblessness is the important factor affecting wages. Hence, while it is not obvious which hypothesis should be preferred, the key feature of this analysis is that unemployment does not outperform either employment density or market potential, and thus the wage curve, by itself, should not be regarded as an outright ‘law’.

Table 7.a. Spatial J-test results: Wage Curve as maintained hypothesis (*dependent variable*: *LogWage*)

	1	2	3	4
	H0:*LogU* H1: UE Model	**H0: *LogU* H1: NEG Model**	**H0:*W*LogU* H1: UE Model**	**H0: *W*LogU* H1: NEG Model**
Endogenous Spatial Lag				
$W*LogWage_{it}\ (\rho)$	-0.0328	0.0047	0.0334	0.0298
(t-stat)	(-1.53)*	(0.22)	(1.63)*	(1.31)
Local Unemployment				
$LogU_{it}$	**-0.0710**	**-0.0246**		
(t-stat)	(-5.34)***	(-2.10)***		
Unemployment within Commuting Distance				
$W*LogU_{it}$			**-0.1203**	**-0.1219**
(t-stat)			(-5.55)***	(-5.67)***
Fitted Values (FV) from UE Model				
FV of $LogWage_{\mathrm{UE}}\ (\delta)$	**1.4684**		**0.8152**	
(t-stat)	(7.40)***		(4.68)***	
Fitted Values (FV) from NEG Model				
FV of $LogWage_{\mathrm{NEG}}\ (\delta)$		**0.9873**		**0.8939**
(t-stat)		(4.95)***		(4.06)***
Local Knowledge Base				
$LogT_{it}\ (\beta_2)$	-0.0156	0.0029	0.0012	-0.0000
(t-stat)	(-1.68)**	(0.30)	(0.14)	(-0.01)
Local Unskilled Workforce				
$LogS_i\ (\beta_3)$	0.1945	0.0443	-0.0026	-0.0000
(t-stat)	(4.01)***	(1.11)	(-0.08)	(-0.01)
Constant	-3.3792	-0.0634	1.0234	0.5489
(t-stat)	(-2.61)***	(-0.05)	(0.92)	(0.39)
Spatial J-Test				
$\hat{J} = \hat{\delta}^2 / \hat{\mathbf{V}}(\delta) \rightarrow^d \chi_1^2$	54.81	24.53	21.90	16.47
Prob > χ_1^2	0.0000	0.0000	0.0000	0.0000
RSS	79.78	85.23	77.40	75.53
$R^{2\,a}$	0.7434	0.7058	0.7510	0.7430
No. Areas	408	408	408	408
No. in-sample years (1999-2009)	11	11	11	11

[a] *Correlation between observed and fitted values of LogWage.*

Table 7.b. Spatial J-test results: UE or NEG as maintained hypotheses (*dependent variable*: *LogWage*)

	1	2	3	4
	H0: UE Model H1: *LogU*	**H0: NEG Model H1: *LogU***	**H0: UE Model H1: *W*LogU***	**H0: NEG Model H1: *W*LogU***
Endogenous Spatial Lag				
$W * LogWage_{it}$ (ρ)	-0.1616	-0.0076	0.0173	0.0039
(t-stat)	(-3.67)***	(-0.18)	(0.93)	(0.21)
Employment Density				
$(LogE_{it} - \rho W * LogE_{it})$	**0.0350**		**0.0214**	
(t-stat)	(8.75)***		(5.98)***	
Market Potential				
$(LogMP_{it} - \rho W * LogMP_{it})$		**0.1211**		**0.1115**
(t-stat)		(5.82)***		(5.35)***
Fitted Values (FV) from Wage Curve Model				
FV of $LogWage_{LogU}$ (δ)	**3.2247**	**1.1168**		
(t-stat)	(6.53)***	(2.49)***		
Fitted Values (FV) from Wage Curve Model				
FV of $LogWage_{W*LogU}$ (δ)			**1.0093**	**0.9837**
(t-stat)			(7.79)***	(7.26)***
Local Knowledge Base				
$LogT_{it}$ (β_2)	-0.1181	-0.0087	-0.0113	-0.0012
(t-stat)	(-5.12)***	(-0.44)	(-1.66)*	(-0.18)
Local Unskilled Workforce				
$LogS_i$ (β_3)	0.1997	0.0178	-0.0477	-0.0000
(t-stat)	(3.18)***	(0.30)	(-1.64)*	(-0.01)
Constant	-14.2787	-1.6337	-0.1824	-0.7102
(t-stat)	(-4.56)	(-0.58)	(-0.22)	(-0.83)
Spatial J-Test				
$\hat{J} = \hat{\delta}^2 / \hat{\mathbf{V}}(\delta) \rightarrow^d \chi_1^2$	42.64	6.21	60.62	52.70
Prob > χ_1^2	0.0000	0.0127	0.0000	0.0000
RSS	78.05	84.83	75.08	75.20
$R^{2\,a}$	0.7431	0.7061	0.7507	0.7424
No. Areas	408	408	408	408
No. in-sample years (1999-2009)	11	11	11	11

[a] *Correlation between observed and fitted values of LogWage.*

8. CONCLUSION

Looking at Britain's 408 local authorities over the period 1998-2010, the present paper has explored the predictive force of the wage curve relatively to two major alternative explanations for spatial wage differentials. These are provided by Urban Economics (UE) and New Economic Geography (NEG) respectively; NEG theory postulates that the level of pay in a given region depends directly and with a distance decay effect on its location relative to large supplier and consumer markets, while UE theory predicts that earnings are higher in cities and larger towns because of external economies and productivity advantages stemming from the spatial concentration of workers and firms. To establish whether the wage curve truly represents an empirical reality, this paper has evaluated its explanatory performance under the direct challenge of each of these competing wage functions considered in turn.

Initially we have estimated an inclusive wage equation that nests both, to see which of the rival models (if any) encompasses the other; we find that unemployment is not dominated by either of the rivals, as it retains its significance in the presence of the competing variable, but there is no evidence that the wage curve is the dominant paradigm, since excluding UE or NEG effects also entails a significant loss of information. To formally discriminate among non-nested models, we have then implemented a spatial J test, alternating between the wage curve and either UE or NEG as to which is treated as the maintained hypothesis. Again, our findings are not conclusive; they show that the wage curve is capable of falsifying either UE or NEG, since the J-statistic rejects the null that its predictive value adds no significance to the maintained (UE or NEG) model, but test results also suggest that we cannot accept the wage curve as the superior proposition in the face of either alternative model.

All in all, our analysis has confirmed that the wage curve holds, as indicated by an estimated slope which is statistically relevant and broadly in line with the classic elasticity of -0.10; additionally, by means of spatial econometrics methods, we have been able to infer that it is unemployment within commuting distance which exerts most influence on wages, with local unemployment accounting for only a small fraction of the total impact. Importantly, however, Blanchflower and Oswald's relationship does not emerge as the only explanation of wage variation, given that UE and NEG are equally successful. The evidence we have presented thus points to the conclusion that the wage curve should not be taken as an absolute principle governing the spatial distribution of wages and economic development; there are in fact other strands of the regional economic literature which are able to account for local wage variation, and these are strongly grounded in economic theory as well as being validated empirically.

REFERENCES

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

Amaral P.V., Lemos M., Simões R., Chein F., 2010, "Regional imbalances and market potential in Brazil", *Spatial Economic Analysis*, 5(4), pp. 463-482.

Baltagi B.H., Blien U., 1998, "The German Wage Curve: Evidence from the IAB Employment Sample", *Economics Letters*, 61, pp. 135-142.

Baltagi B.H., Blien U., Wolf K., 2000, "The East German wage curve: 1993-1998", *Economics Letters*, 69, pp. 25-31.

Baltagi B.H., Blien U., Wolf K., 2009, "New Evidence on the Dynamic Wage Curve for Western Germany: 1980–2004", *Labour Economics*, 16, pp. 47-51.

Barde S., 2010, "Increasing returns and the spatial structure of French wages", *Spatial Economic Analysis*, 5, pp. 73-91.

Bell B., Nickell S., Quintini G., 2002, "Wage Equations, Wage Curves and All That", *Labour Economics*, 9, pp. 341-360.

Blanchflower D.G., Oswald A.J., 1990, "The Wage Curve", *Scandinavian Journal of Economics*, 92, pp. 215-235.

Blanchflower D.G., Oswald A.J., 1994a, *The Wage Curve*, Cambridge: MIT Press.

Blanchflower D.G., Oswald A.J., 1994b, "Estimating a Wage Curve for Britain 1973-90", *The Economic Journal*, 104, pp. 1025-1043.

Blanchflower D.G., Oswald A.J., 1995, "An Introduction to the Wage Curve", *Journal of Economic Perspectives*, 9, pp. 153-167.

Blanchflower D.G., Oswald A.J., 2005, "The Wage Curve Reloaded", Working Paper No. 11338, *National Bureau of Economic Research*, Cambridge, MA.

Brakman S., Garretsen H., Schramm M., 2004, "The spatial distribution of wages and employment: estimating the Helpman-Hanson model for Germany", *Journal of Regional Science*, 44(3), pp. 437-466.

Brakman S., Garretsen H., Van Marrewijk C., 2009a, *An Introduction to Geographical Economics*, Cambridge: Cambridge University Press.

Brakman S., Garretsen H., Van Marrewijk C., 2009b, "Economic geography within and between European Nations: the role of market potential and density across space and time", *Journal of Regional Science*, 49, pp. 777–800.

Brülhart M., Mathys N.A., 2008, "Sectoral agglomeration economies in a panel of European regions", *Regional Science and Urban Economics*, 38, pp. 348-362.

Buettner T., 1999, "The Effect of Unemployment, Aggregate Wages, and Spatial Contiguity on Local Wages: an Investigation with German District Level Data", *Papers in Regional Science*, 78, pp. 47-67.

Burridge P., Fingleton B., 2010, "Bootstrap inference in spatial econometrics: the J test", *Spatial Economic Analysis*, 5(1), pp. 93-119.

Card D., 1995, "The Wage Curve: A Review", *Journal of Economic Literature*, 33, pp. 785-799.

Ciccone A., Hall R.E., 1996, “Productivity and the density of economic activity”, *American Economic Review*, 86, pp. 54-70.

Combes P-P., Duranton G., Gobillon L., 2008, “Spatial wage disparities: sorting matters!”, *Journal of Urban Economics*, 63, pp. 723-742.

Combes P-P., Mayer T.H., Thisse J-F, 2008, *Economic Geography: the Integration of Regions and Nations*, Princeton: Princeton University Press.

Corrado L., Fingleton B., 2012, “Where is the economics in spatial econometrics?”, *Journal of Regional Science*, 52, pp. 210-239.

Davidson R., MacKinnon J., 1981, “Several tests for model specification in the presence of alternative hypotheses”, *Econometrica*, 49, pp. 781-93.

Davidson R., MacKinnon J., 1982, “Some non-nested hypothesis tests and the relations among them”, *Review of Economic Studies*, 49, pp. 551-65.

Davidson R., MacKinnon J., 1993, *Estimation and Inference in Econometrics*, Oxford: Oxford University Press.

Davidson R., MacKinnon J., 2002a, “Bootstrap J tests of non-nested linear regression models”, *Journal of Econometrics*, 109, pp. 167-93.

Davidson R., MacKinnon J., 2002b, “Fast double bootstrap tests of nonnested linear regression models”, *Econometric Reviews*, 21, pp. 419-29.

Dixit A., Stiglitz J.E., 1977, “Monopolistic competition and optimum product diversity”, *American Economic Review*, 67, pp. 297-308.

Elhorst J.P., Blien U., Wolf K., 2007, “New Evidence on the Wage Curve: A Spatial Panel Approach”, *International Regional Science Review*, 30(2), pp. 173-191.

Elhorst J.P., 2010a, “Spatial Panel Data Models”, in Fischer M.M., Getis A. (eds) *Handbook of Applied Spatial Analysis,* Berlin, Germany: Springer.

Elhorst J.P., 2010b, “Applied spatial econometrics: raising the bar”, *Spatial Economic Analysis*, 5, pp. 9-28.

Fingleton, B., 2003, “Increasing returns: evidence from local wage rates in Great Britain”, *Oxford Economic Papers*, 55, pp. 716-39.

Fingleton B., 2006, “The new economic geography versus urban economics: an evaluation using local wage rates in Great Britain”, *Oxford Economic Papers*, 58, pp. 501-30.

Fingleton B., 2007, “Testing the ‘new economic geography’: a comparative analysis based on EU regional data”, in Fingleton B. (ed.) *New Directions in Economic Geography*, Cheltenham: Edward Elgar, pp. 70-97.

Fingleton B., 2008, “Competing models of global dynamics: evidence from panel models with spatially correlated error components”, *Economic Modelling*, 25(3), pp. 542-558.

Fingleton B., Longhi S., 2013, “The effects of agglomeration on wages: evidence from the micro-level”, *Journal of Regional Science*, 53(3), pp. 443-463.

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

Fujita M., Krugman P.R., Venables A.J., 1999, *The Spatial Economy: Cities, Regions, and International Trade*, Cambridge, MA: MIT Press.

Greene W.H., 2003, *Econometric Analysis*, 5th edition, New Jersey: Prentice Hall.

Hanson G., 1998, "Market potential, increasing returns, and geographic concentration", *Journal of International Economics*, 67(1), pp. 1-24.

Harris C., 1954, "The market as a factor in the localization of industry in the United States", *Annals of The Association of American Geographers*, 64, pp. 315-348.

Head K., Mayer T., 2004, "Market potential and the location of Japanese investment in the European Union", *Review of Economics and Statistics*, 86(4), pp. 959-972.

Hendry D.F., 1995, *Dynamic Econometrics*, Oxford: Oxford University Press.

Huriot J-M., Thisse J-F., 2000, *Economics of Cities*, Cambridge: Cambridge University Press.

Jacobs J., 1969, *The Economy of Cities*, New York: Vintage Books.

Johnes G., 2007, "The Wage Curve Revisited: Estimates from a UK panel", *Economics Letters*, 94, pp. 414-420.

Kapoor M.H., Kelejian H.H., Prucha I.R., 2007, "Panel data models with spatially correlated error components", *Journal of Econometrics*, 140, pp. 97-130.

Kelejian H.H., 2008, "A spatial J-test for model specification against a single or a set of non-nested alternatives", *Letters in Spatial and Resource Sciences*, 1, pp. 3-11.

Kelejian H.H., Prucha I.R., 1998, "A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances", *Journal of Real Estate Finance and Economics*, 17, p. 99-121.

Kennedy S., Borland J., 2000, "A wage curve for Australia?", *Oxford Economic Papers*, 52, pp. 774-803.

Krugman P.R., 1991, "Increasing returns and economic geography", *Journal of Political Economy*, 99, pp. 483-499.

Krugman P.R., Venables A.J., 1995, "Globalization and the inequality of nations", *Quarterly Journal of Economics*, 110, pp. 857-80.

Le Sage J., Pace R.K., 2009, *Introduction to Spatial Econometrics*, Boca Raton, FL: CRC Press.

Longhi S., Nijkamp P., Poot J., 2006, "Spatial Heterogeneity and the Wage Curve Revisited", *Journal of Regional Science*, 46, pp. 707-731.

Manning N., 1994, "Earnings, unemployment and contiguity: evidence from British counties 1976-1992", *Scottish Journal of Political Economy*, 41(1), pp. 43-68.

Mion G., 2004, "Spatial externalities and empirical analysis: the case of Italy", *Journal of Urban Economics*, 56(1), pp. 97-118.

Mion G., Naticchioni P., 2009, "The spatial sorting and matching of skills and firms", *Canadian Journal of Economics*, 42(1), pp. 28-55.

Molho I., 1995, "Spatial autocorrelation in British unemployment", *Journal of Regional Science*, 35(4), pp. 641-658.

Morrison P.S., Papps K.L., Poot, J., 2006, "Wages, employment, labour turnover and the accessibility of local labour markets", *Labour Economics*, 13, pp. 639-663.

Moulton B.R., 1990, "An illustration of a pitfall in estimating the effects of aggregate variables on micro units", *Review of Economics and Statistics*, 72(2), p. 334-338.

Niebuhr A., 2006, "Market access and regional disparities", *The Annals of Regional Science*, 40(2), pp. 313-334.

Nijkamp P., Poot J., 2005, "The Last Word on the Wage Curve ?", *Journal of Economic Surveys*, 19, pp. 421-450.

Pesaran M.H., 2006, "Estimation and inference in large heterogenous panels with a multifactor error structure", *Econometrica*, 74, pp. 967-1012.

Rivera-Batiz F., 1988, "Increasing returns, monopolistic competition, and agglomeration economies in consumption and production", *Regional Science and Urban Economics*, 18, pp. 125-153.

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

Sato Y., 2000, "Search theory and the wage curve", *Economics Letters*, 66, pp. 93-98.

Venables A.J., 1996, "Equilibrium locations of vertically linked industries", *International Economic Review*, 37, pp. 341-59.

APPENDIX. Derivation of the Urban Economics Model

The starting point is a *Cobb-Douglas production function for the output Q* of the competitive final goods and services sector: $Q = (M^{\beta} I^{1-\beta})^{\alpha}$. Internal increasing returns are modelled through the monopolistic competition and product variety theory of Dixit and Stiglitz (1977), assuming a *CES* (Constant Elasticity of Substitution) *sub-production function for producer service inputs I* (the increasing-returns sector[13]). This means that intermediate services are modelled as a 'continuum' of *x* varieties, each produced by a specialised firm with monopolistic power[14], with *i(t)* representing the amount of type-*z* variety in the assumed 'continuum' of varieties.

$$I = \left[\sum_{z=0}^{x} i_z^{(\sigma-1)/\sigma} \right]^{\sigma/(\sigma-1)} = \left[\sum_{z=0}^{x} i_z^{1/\mu} \right]^{\mu}$$

[13] *I*-sector activities include business and professional services, financial services, insurance services, and real estate services. These sub-sectors can be considered as being characterized by small firms producing highly-differentiated varieties, easy entry and exit, and minimal strategic interaction, which is close to what is implied by monopolistic competition.

[14] Monopolistic power exist because internal scale economies imply large-size firms with market power, and thus debar perfect competition, and also because each firm produces a differentiated variety of the composite intermediate service.

In equilibrium, the CES sub-production function can be re-written as $I = x^{\mu} i$ because, due to the assumption of free entry and exit in response to positive and negative profits, each firm produces the same zero-profit level of intermediate service output, equal to *i* for all *z*. In $I = x^{\mu} i$, parameter **μ (μ>1)** measures the equilibrium amount of *internal increasing returns to scale* that can be exploited by the individual producer service firm, since an increase in the number of firms/varieties *x* yields a more than proportionate increase in the output of the intermediate service sector *I*. It also determines the *constant price elasticity of demand* (*ped*), since from the constant elasticity demand function

$$i(z) = kp^{-\sigma} = kp^{-\mu/(\mu-1)}$$

$$\mathrm{d}i(z)/\mathrm{d}p = k \cdot -\frac{\mu}{\mu-1} p^{-\mu/(\mu-1)-1} = k \cdot -\frac{\mu}{\mu-1} p^{-\mu/(\mu-1)} p^{-1} = -\frac{kp^{-\mu/(\mu-1)}\mu}{(\mu-1)p} = -\frac{i(z)\mu}{(\mu-1)p}$$

$$ped = -\frac{\mathrm{d}i(z)}{\mathrm{d}p} \cdot \frac{p}{i(z)} = -\left(-\frac{i(z)\mu}{(\mu-1)p}\right) \cdot \frac{p}{i(z)} = \frac{i(z)\mu}{(\mu-1)p} \cdot \frac{p}{i(z)} = \frac{\mu}{\mu-1}$$

It also determines the *constant elasticity of substitution* among varieties (σ), since

$$-\frac{di(z)}{dp_z} \frac{p_z}{i(z)} = \frac{\mu}{\mu-1} = \sigma$$

Hence μ controls the *degree of differentiation in the service sector* (i.e. the importance of intermediate service variety to final producers), because as μ shrinks to one the elasticity of substitution σ increases, and also the *degree of monopolistic/market power* available to producer service firms, which decreases as varieties become more perfectly substitutable.

Production costs (i.e. labour costs, as labour is the only input) incurred by each producer service firm have a fixed component as well as a variable component; there is a fixed labour requirement to start production, *s*, then labour requirement increases with output, i.e. $a \cdot i(z)$ with *a* being the marginal labour requirement. The presence of fixed production costs implies falling average costs as output increases, so that a firm can obtain a cost advantage and operate more efficiently by concentrating production at one large plant. Therefore, the amount of labour efficiency units required to produce each variety is given by

$$\left[a \cdot i(z) + s\right]$$

The number of producer service firms/varieties *x* is equal to producer service effective labour in the unit area divided by effective labour per firm

$$x = \frac{(1-\beta) \cdot N}{a \cdot i(z) + s}$$

where (1-β) is the share of total labour efficiency units employed in the producer service sector under competitive equilibrium in the labour market. Thus an increase in density of activity (in the form of employment density *N*) leads to an increase in the variety of intermediate producer services *x*, because a larger number of firms can break even when the local market is larger; however, the scale of production of any one existing variety remains unchanged, with each of the x firms producing the same zero-profit level of output $i(z)$. Moreover, since $I=x^{\mu}i$ and since it requires $x{\cdot}i$ units to produce I, intermediate service productivity is $x^{\mu-1}$, so that there is a positive relation between the number of producer service firms/varieties x and productivity $x^{\mu-1}$.

Replacing I with $x^{\mu}i$ in $Q=(M^{\beta}I^{1-\beta})^{\alpha}$ gives

$$Q=(M^{\beta}(x^{\mu}i(z))^{1-\beta})^{\alpha}=M^{\alpha\beta}x^{\alpha\mu(1-\beta)}i(z)^{\alpha(1-\beta)}$$

and substituting for $M=\beta N$ and for $x=\dfrac{(1-\beta)\cdot N}{a\cdot i(z)+s}$ we obtain

$$\begin{aligned}
Q&=(\beta N)^{\alpha\beta}\left(\frac{(1-\beta)N}{a\cdot i(z)+s}\right)^{\alpha\mu(1-\beta)}i(z)^{\alpha(1-\beta)}\\
&=\beta^{\alpha\beta}N^{\alpha\beta}N^{\alpha\mu(1-\beta)}(1-\beta)^{\alpha\mu(1-\beta)}\left(a\cdot i(z)+s\right)^{-\alpha\mu(1-\beta)}i(z)^{\alpha(1-\beta)}\\
&=\beta^{\alpha\beta}N^{\alpha\beta}N^{-\alpha\mu(\beta-1)}(1-\beta)^{-\alpha\mu(\beta-1)}\left(a\cdot i(z)+s\right)^{\alpha\mu(\beta-1)}i(z)^{\alpha(1-\beta)}\\
&=N^{\alpha(\beta+\mu-\mu\beta)}\beta^{\alpha\beta}(1-\beta)^{-\alpha\mu(\beta-1)}\left(a\cdot i(z)+s\right)^{\alpha\mu(\beta-1)}i(z)^{\alpha(1-\beta)}
\end{aligned}$$

Collecting constants simplifies to

$$Q=\phi N^{\alpha(1+(1-\beta)(\mu-1))}=\phi N^{\gamma}$$

LA COURBE DE SALAIRE RECONSIDÉRÉE : S'AGIT-IL VRAIMENT D'UNE LOI EMPIRIQUE DE L'ÉCONOMIE ?

Résumé *- La relation négative entre les salaires et le chômage, appelée «courbe de salaire », a fait l'objet d'une vaste littérature et a été définie comme une « loi empirique de l'économie ». Cependant, des théories alternatives plus récentes, issues de l'Economie Urbaine et de la Nouvelle Economie Géographique, tentent d'expliquer la variation des salaires sans faire référence au taux de chômage. Cet article étudie la représentativité des modèles non imbriqués de détermination des salaires en utilisant des données concernant 408 communes britanniques (local authorities) sur la période 1998-2010.*

Mots-clés : COURBE DE SALAIRE, DENSITÉ D'EMPLOI, MARCHÉ POTENTIEL, ÉCONOMÉTRIE SPATIALE DES DONNÉES DE PANEL, J-TEST SPATIAL, MODÈLES NON IMBRIQUÉS