University of Strathclyde

Department of Economics

# Essays in Spatial Panel Econometrics

by

Silvia Palombi

A thesis presented in fulfilment of the requirements for the degree of Doctor of Philosophy

2015

# Declaration of Authenticity and Author's Rights

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.



Signed: ………………………………………………….
Silvia Palombi

Date: ….………………………………………………….

# Statement of Contribution of Co-Authors

Chapters 2 to 5 consist of previously published work or papers submitted for publication and under editorial review during the period of candidature. The undersigned hereby certify that:

1. they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
2. they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
3. there are no other authors of the publication according to these criteria; and
4. potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit.

In the case of Chapter 2, contributions to the work involved the following:

| **Contributor** | **Statement of contribution** | **Publication title and date of publication or status** |
|---|---|---|
| B. Fingleton<br><br>Signature | FGS2SLS/GMM econometric routines<br>Methodological guidance<br>Wrote manuscript (introduction, data analysis) | Fingleton, B. and Palombi, S. (2013) Spatial panel data estimation, counterfactual predictions, and local economic resilience among British towns in the Victorian era, *Regional Science and Urban Economics*, 43(4), 649-660. |
| S. Palombi<br><br>Signature | Data preparation<br>Literature review<br>Implemented fitting and forecasting procedure<br>Documented empirical results<br>Wrote manuscript (estimation method, data analysis, concluding remarks) | Fingleton, B. and Palombi, S. (2013) Spatial panel data estimation, counterfactual predictions, and local economic resilience among British towns in the Victorian era, *Regional Science and Urban Economics*, 43(4), 649-660. |

In the case of Chapter 3, contributions to the work involved the following:

| Contributor | Statement of contribution | Publication title and date of publication or status |
|---|---|---|
| B. Fingleton<br><br>Signature | Published previously on non-nested wage theories for the UE vs NEG case<br>Supervised work<br>Checked manuscript | Fingleton, B., and Palombi, S. (2013) The Wage Curve reconsidered: is it truly an empirical law of economics?, *Région et Développement*, 38, 49-92. |
| S. Palombi<br><br>Signature | Data preparation<br>Literature review<br>Wrote wage curve model specification<br>Wrote model comparison and selection methodology<br>Conducted empirical estimation<br>Wrote manuscript | Fingleton, B., and Palombi, S. (2013) The Wage Curve reconsidered: is it truly an empirical law of economics?, *Région et Développement*, 38, 49-92. |

In the case of Chapter 4, contributions to the work involved the following:

| Contributor | Statement of contribution | Publication title and date of publication or status |
|---|---|---|
| B. Fingleton<br><br>Signature | Econometric routines<br>Supervised work<br>Checked manuscript | Bootstrap *J*-test for panel data models with spatially dependent error components, a spatial lag and additional endogenous variables, accepted for publication in *Spatial Economic Analysis*, DOI: 10.1080/17421772.2016.1102960. |
| S. Palombi<br><br>Signature | Literature review<br>Wrote *J*-test specification<br>Conducted bootstrap simulations<br>Set up & ran Monte Carlo trials<br>Wrote manuscript | Bootstrap *J*-test for panel data models with spatially dependent error components, a spatial lag and additional endogenous variables, accepted for publication in *Spatial Economic Analysis*, DOI: 10.1080/17421772.2016.1102960. |

In the case of Chapter 5, contributions to the work involved the following:

| **Contributor** | **Statement of contribution** | **Publication title and date of publication or status** |
|---|---|---|
| S. Palombi<br><br>Signature | Literature review<br>Data analysis<br>Wrote manuscript | Commuting effects in Okun's Law among British areas: evidence from spatial panel econometrics, accepted for publication in *Papers in Regional Science*, DOI: 10.1111/pirs.12166. |
| R. Perman<br><br>Signature | Checked manuscript and provided feedback | Commuting effects in Okun's Law among British areas: evidence from spatial panel econometrics, accepted for publication in *Papers in Regional Science*, DOI: 10.1111/pirs.12166. |
| C. Tavéra<br><br>Signature | Checked manuscript and provided feedback | Commuting effects in Okun's Law among British areas: evidence from spatial panel econometrics, accepted for publication in *Papers in Regional Science*, DOI: 10.1111/pirs.12166. |

# Acknowledgements

I would like to thank my advisors, Bernard Fingleton and Roger Perman, for all their support and guidance through my PhD years. I would also like to thank everyone in the Economics Department for their encouragement and for helpful and stimulating discussions.

Many thanks go to colleagues at the Office of the Chief Economic Adviser, who inspired me and help me stay focused and motivated; their encouragement while I was completing my PhD after joining the Scottish Government has been invaluable. I am also grateful to family and friends who kept me going throughout the last years.

Finally, this thesis was supported financially by the Department of Economics of the University of Strathclyde which I gratefully acknowledge.

# Abstract

This thesis develops and applies state-of-the-art spatial panel econometrics methods in order to model and analyse labour-market outcomes within and across small areas in Great Britain, with respect to three particular aspects; namely the resilience of local economies to periodic shocks, the determinants of spatial disparities in local wages, and the relationship between output and unemployment over the economic cycle. The contribution is thus provided in four essays.

The first essay (Chapter 2) explores the relative ability of local economies to preserve their long-run growth dynamics when faced by the destabilising effects of major shocks. It borrows concepts from the regional economic resilience literature to characterise the different reactions of different places to recessions. In particular, economies are distinguished on the basis of their ability to resist to and recover from shocks thus maintaining stability around their counterfactuals, a notion known as 'engineering resilience', and to resume (or improve) their underlying growth trajectory by the end of the recessionary period under consideration thus showing 'ecological resilience' (or 'positive hysteretic effects'). Related to these notions are the ideas of adaptability and 'path dependence', which help explain why some economies are more vulnerable to shocks than others (over and above the static causes of interregional heterogeneity incorporated in the model via random effects). Taking annual wage series for nineteen British towns over the historical period 1871-1906, I fit a spatial panel data model to 1871-1890 data by Spatial Two-Stage Least Square / Generalised Method of Moments (S2SLS / GMM), and use estimated coefficients in combination with trend forecasts to obtain counterfactual predictions of wage levels after the 1890 shock through to 1906. This allows to analyse how actual wages in different towns performed in relation to their counterfactual paths, and to gauge their relative resilience to economic shocks. The key finding, and the main lesson that can be drawn from the historical experience of British towns, is that the sectoral composition of local employment is important for economic resilience;

my evidence suggests that excessive and increasing specialisation in declining industries means lack of the structural flexibility needed to replace these industries with competitive and productive activities (shock-proneness), whereas economies with a diversified industrial mix have greater scope for restructuring and renewal (shock-resilience); moreover, towns dominated by mature, staple sectors but who have also developed new growth industries are more able to adapt to and tolerate shocks.

The second essay (Chapter 3) considers the relative success of alternative, non-nested wage equations from the perspective of Great Britain's 408 unitary authority and local authority districts (UALADs) over the period 1999-2009. 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 theories that seek to explain regional wage variations without reference to unemployment, namely Urban Economics (UE) and New Economic Geography (NEG). The aim is to discriminate between competing models of wage determination in order to establish whether the wage curve can be accepted as superior to its non-nested rivals. To do so I adopt an 'Inclusive Regression' approach (Davidson and MacKinnon, 1993; Hendry, 1995), combining the wage curve and either UE or NEG within an Artificial Nesting Model (ANM); this incorporates a spatial autoregressive process involving both the dependent variable and the error components and is estimated by S2SLS / GMM. The main conclusion is that, at least when the level of geographical resolution is relatively low as in this sample, while being validated empirically the wage curve should not be taken as an outright 'law' governing the spatial wage distribution. Specifically, when using asymptotic *P*-values, the wage curve is not dominated by either UE or NEG, in the sense that unemployment retains its predictive power in the presence of either employment density or market potential, nor is it capable of falsifying its rivals, as each of these is also statistically significant under the ANM. Meanwhile, when using bootstrap *P*-values, the wage curve emerges as the leading statement when directly confronted by NEG whereas UE offers a seemingly adequate hypothesis.

The third essay (Chapter 4) is an extension of the previous chapter, and provides conclusive evidence by implementing a more formal and rigorous approach to testing a null model against a non-nested alternative, i.e. the *J*-test. This is a well-established technique for choosing among non-nested rivals, and in this chapter I develop a version of the test for specifications (SARAR-RE models) which feature spatially correlated error components, thus accounting for interregional heterogeneity via random effects (also subjected, like the disturbances, to a spatially autoregressive process), as well as a spatial lag of the dependent variable and additional, potentially endogenous regressors. This chapter thus makes a valuable addition to the literature on non-nested hypotheses testing in the spatial panel context by extending the toolkit to random-effects models. I also provide Monte Carlo evidence showing that there are distributional issues associated with the asymptotic use of the *J*-test in small-to-medium samples, so another novelty of this chapter is the implementation of a Bootstrap scheme to construct a valid null reference distribution in finite samples when the null and alternative are SARAR-RE models estimated by S2SLS / GMM. In terms of the empirical application, bootstrap *J*-test results confirm the bootstrap ANM results from the previous chapter that the wage curve rejects NEG theory while UE theory is equally successful. Another finding, from the methodological angle, is that the bootstrap *J*-test is a reliable and effective procedure for correcting asymptotic reference critical values and distinguishing between competing hypotheses in all cases where one is not a reduced form of the other.

The fourth and final essay (Chapter 5) is one of few to reconsider from a spatial panel econometric perspective an economic relationship - the ‘empirical law of economics’ known as Okun’s Law - which has been traditionally considered at macro level with no attention for sub-national phenomena; it is the first to do so for Great Britain, looking at the 128 British NUTS3 regions over the period 1985-2011. By means of specialist techniques recently devised for spatial data, I show that regional interdependencies have a prominent role in the unemployment-output relationship; the total Okun’s Law effect itself is close to the ‘law’ of -0.30 but more than two thirds of this are accounted for by the impact on local unemployment rate of real output variations in areas nearby, a finding suggesting that policy intervention at both national and regional level on a country’s labour market can be more effective if spatial effects are factored into the analysis and modelled / tested explicitly.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Overview

This thesis, comprising four separate but interrelated chapters, develops and applies state-of-the-art spatial panel econometric methods to specific research problems in regional economics. All chapters are based on statistical information for Great Britain but the datasets cover different time periods and geographical units of different type. Moreover, as evident below, each chapter attempts to make both a methodological and an empirical contribution.

Chapter 2 focuses on estimation and prediction with a spatial panel data model, in the context of the literature on regional disparities in economic resilience to recessionary shocks, using wage data for a sample of British towns in the Victorian era. Chapter 3 is concerned with comparison and selection among non-nested competing hypotheses of spatial wage imbalances, looking at recent data for the British UALADs; it presents initial findings on the performance of the Wage Curve (commonly regarded to as an ‘empirical law of economics’) relative to its Urban Economics (UE) and New Economic Geography (NEG) rivals using an ‘Inclusive Regression’ or ‘Artificial Nesting Model’ approach. Chapter 4 is an extension of Chapter 3 which provides conclusive evidence on the research problem in the third chapter; it develops and implements a random-effects version of the spatial *J*-test procedure for testing a null model against a non-nested alternative, using Bootstrap techniques to construct a valid finite-sample reference distribution for the test statistic under the null. Chapter 5 continues on the subject of spatial modelling, revisiting Okun’s Law (also referred to as an ‘empirical law of economics’) using panel time-series data for the British NUTS3 regions; it fits an unemployment-output relationship augmented with spatial effects, with an attention to the correct estimation and interpretation of model parameters in cases where the specification incorporates spatial lags of the dependent and independent variables.

From an empirical perspective, the thesis is constructed so as to answer the following research questions. First (Chapter 2), can a spatial panel data model be used to generate counterfactual forecasts of local wages, and to cast light on the possible reasons why the labour market in some places are more resilient than others to recessionary shocks? Second (Chapters 3 and 4), is the Wage Curve empirically validated by the spatial panel dataset under study, and is it truly an 'empirical law of economics' or the newer NEG and UE theories perfom better at explaining the space-time distribution of local wages and should thus be used to understand the geography of wages? Finally (Chapter 5), does Okun's Law hold given the spatial panel time-series dataset at hand and can something be learnt from the inclusion of spatial effects about how demand policies feed through to labour-market outcomes?

Therefore, with the increasing availability of spatial panel data calling for new modelling procedures, this thesis has endeavoured to equip the applied researcher with econometric techniques for estimation, prediction and inference with such data, and has looked at important issues in empirical economics such as economic resilience, economic agglomeration and unemployment to show how these methods can be used in practice.

## 1.2 Methodological background

An issue of prime interest, when using panel data, is whether to opt for fixed effects or random effects. Below I summarise the main reasons for the popularity of random-effects models, as extensively discussed in the available literature (e.g., most recently, Baltagi, 2013; Elhorst, 2014, section 3.4), and their merits in the context of this thesis.

- A fixed-effects model is particularly indicated when regression analysis involves a **precise set of individual units convering the whole population**, while a random-effects model is a more appropriate specification if a certain number of individuals are drawn at random from a larger population (Arbia, Basile and Piras, 2005). This is one of the reasons for adopting random effects in Chapter 2, which uses a random sample of nineteen towns.

The data in Chapters 3 and 4 and in Chapter 5 do not comprise a sample, however it is possible to consider the data in these cases to also be one of many realisations from a superpopulation since the spatial partitions giving the areal units are just **one of an infinite number of possible sets** that could have occurred (see also Fingleton, 2010, p. 5, note 12). Thus, also in Chapters 3 to 5, I prefer not to condition inference on the spatial units of observation within the study area, but to try to relate to a larger (hypothetical) population.

- Fixed-effects and random-effects models also differ in that they use different parts of the variation in the data (Partridge, 2005; Baltagi, 2013; Elhorst, 2014, section 3.4). Panel data models with controls for spatial fixed effects utilise solely the time-series dimension of the data, whereas random-effects models exploit **both time-series and cross-sectional information**, thus improving the precision of estimates.
- As a result it is argued that, by taking account of permanent (cross-sectional or between-unit) variation, spatial random effects tend to give **long-run estimates**, whereas within-unit fixed-effects estimation focuses on short-term variation. This is especially important in Chapter 2, where the model must be specified in such way that it is able to pick up hysteretic effects.
- Related to this is the fact that random-effects estimation permits the **identification of covariates with minimal or no time variation**, which would be

either problematic or impossible otherwise. This is relevant to Chapters 3 and 4, where there is a time-invariant variable which would be disallowed under fixed effects.

- Also, this means that when, as in Chapters 3 and 4, $T$ is very small compared to $N$, the **vast bulk of variation in spatial data is between- rather than within-sample**, thus making fixed-effects (within-based) estimation dubious or incorrect. Moreover, with fixed-effects models, because the number of observations available to estimate each $u_i$ is given by $T$, the number of time periods must be sufficiently large to have unbiased individual effects estimates; this does not matter when the slope coefficients estimates are of interest and the fixed effects estimates are not, since slope coefficients are not a function of fixed effects, and thus the problem is not transmitted from the latter to the former.
- Another advantage of random-effects estimation, where any omitted (permanent) causes of individual heterogeneity are modelled as being part of the composite spatial error term, is that the random-effects component is itself allowed to be **potentially subject to spatial correlation**, which can reduce the potential for bias in the estimated standard errors and improve inference about wage determinants.

The decision for Chapters 2 to 4 is therefore to use random effects because fixed effects are not a valid option, for the different reasons that are highlighted in the above discussion. Specifically, Chapter 2 uses a random sample of nineteen towns, rather than a precise set of individual areas covering the whole population, moreover it is important that the econometric model defining the impact and transmission of shocks must be specified in such way that it is able to pick up long-run effects. For Chapters 3 and 4, the main motivation for the choice of random effects is that the model includes a time-invariant variable, which would not be possible to identify under fixed effects.

In Chapter 5, differently from Chapters 2 to 4, I estimate spatial panel data models with either fixed effects or random effects to control for cross-region heterogeneity. The justification is that I am less certain about the appropriateness of random effects in this case, as this is not concerned with a random sample as the first chapter, there are no time-constant covariates, and the time dimension ($T$=27) is larger and the spatial dimension ($N$=128) smaller than in the second and third chapters ($T$=11 and $N$=408).

This means that a Hausman consistency test, while absent in the early chapters given the required random-effects strategy, is necessary in Chapter 5 in order to compare Okun's Law coefficients under both estimation approaches and choose the method which gives the best estimates.

It is important to clarify the measurement of the unemployment variable, and the interpretation of the regression coefficient of interest, in the context of the Wage Curve (Chapters 2, 3, 4) and in the context of Okun's Law (Chapter 5).

For the Wage Curve, using general notation, I have

$$\ln(wage) = \alpha + \beta \ln(unemployment\ rate) + \varepsilon$$

Unemployment is a rate, and the 'empirical law' is a coefficient $\beta$ of -0.10. This means that a 1% rise in the local unemployment rate, for example from 5% to 5.05%, is associated with a wage reduction of approximately 0.10%, for example from £300 gross per week to £299.70.

For Okun's Law, unemployment is the dependent variable. Using the usual notation, I have

$$Unemp\ rate\ change\ (percentage\ points) = \alpha + \beta\left(GDP\ growth\ rate\ (\%)\right) + \varepsilon$$

The 'empirical law' is a 3:1 trade-off between economic growth and unemployment rate changes. It predicts that a 1% increase in the (real) GDP growth rate, for example from 1% to 2%, yields a 0.3-0.5 percentage point fall in the unemployment rate, for example from 5% to 4.5-4.7%.

# Chapter 2

# Estimation and Prediction with Spatial Panel Data: A Counterfactual Analysis of Local Economic Resilience among British Towns in the Victorian Era

## 2.1 Introduction

This chapter is concerned with the differentiated impact of economic shocks on wage levels across towns in England, Scotland and Wales over the period 1871-1906. It differs from econometric impulse response models[2] in that it is based on spatial econometric models incorporating simultaneous, *global* spillovers across space in a panel data context as well as hysteretic effects. The model allows estimation of the underlying trend in the evolution of wages, and this acts as a counterfactual against which the actual wage series can be compared.

The data relate to the historical Victorian period in which the industrial revolution was at its height. The economy of the time was, as now, subject to major shocks and the analysis in this essay is aimed at exploring whether these shocks had a permanent effect on the subsequent evolution of wages in different towns. In particular, the study focuses on the mid-period shock of 1890[3], estimating the model up to this point in time, and then projecting the underlying trend forward. The predicted wage path for each town, from which the effects of earlier shocks have been purged, acts as a counterfactual. This counterfactual, or projected, path is what

[2] Fingleton, Garretsen and Martin (2012) fit a Vector Error Correction Model (VECM) to levels of employment (in millions) in UK NUTS1 regions over the period 1971-2009, and use Impulse Response Functions (IRFs) to analyse the impact of a one-time unit shock to employment in the South East. They find that the effects are non-temporary, given that they do not die out to zero over time, and that, as well as effects applying to the own region, there are also interregional responses involving nearby regions.

[3] This choice is convenient in that it allows to have time periods of suitable length for both estimation and forecasting.

one would expect wages to follow if the economy was resilient to the 1890 and subsequent shocks.

Towns that failed to recover from the shocks by 1906, and had wages consistently below the counterfactual, are those that mostly felt the negative effects of the economic shocks, which indicates their shock-proneness. Towns where wage levels were close to the projected path in 1906, and which maintained overall stability around the counterfactual, can be regarded to as resilient. Other towns which ended up with wage levels above the projected path, and which performed well relative to the counterfactual with wages higher or similar on average, one might think of as super-resilient.

Next I explore possible reasons for these different responses to major periodic crises, drawing on the literature on economic resilience (Holling, 1973, 1996, 2001; Pim, 1984; McGlade et al., 2006; Walker et al., 2006; Hill et at., 2008), hysteresis (Blanchard and Summers, 1987; Cross and Allen, 1988; Cross, 1993; Cross, Grinfeld and Lamba, 2009; Cross, Mcnamara and Pokrovskii, 2010; Setterfield, 2010), and path dependence and lock-in (Arthur, 1989, 1990, 1994; David, 1985, 2005, 2007; Martin and Sunley, 2006, 2009; Boschma and Martin, 2009; Simmie and Martin, 2010; Martin, 2012).

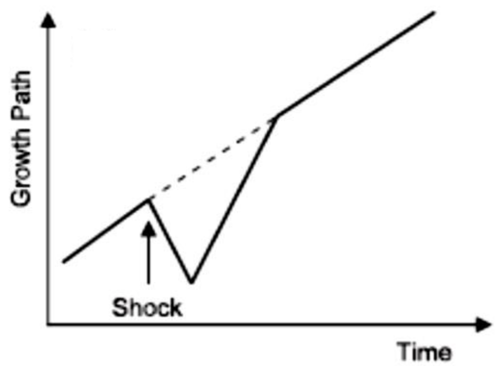
The chapter is structured as follows. Section 2.2 outlines the regional economic resilience framework. Section 2.3 illustrates the model of wage determination adopted to obtain parameter estimates for in-sample data (1872-1890) which are then used to obtain ex-post counterfactual wage predictions for the out-of-sample period (1890-1906). The estimation procedure, including the instrumentation strategy, is shown in section 2.4, while section 2.5 presents the in-sample empirical results. Section 2.6 explains the out-of-sample counterfactual prediction exercise, while results from the counterfactual analysis are discussed in section 2.7. The final section summarises and concludes.

## 2.2 The regional economic resilience framework

A useful framework with which to think about how local areas cope during economic downturns, and to explain why some areas are more sensitive to recessions than others, is provided by the literature on regional economic resilience. Below I review the main conceptual strands, which are at the base of the counterfactual analysis in section 2.7 where I discuss observed patterns and possible causes of spatial disparities in shock impacts.

### *2.2.1 Equilibrium approaches to resilience: 'engineering' perspective*

The 'engineering' version of economic resilience (Holling, 1973, 1996, 2001; Pimm, 1984; Walker et al., 2006) is similar to the 'Plucking Model' idea of economic fluctuations (Friedman, 1993; Kim and Nelson, 1998). The basic assumption is that a local economy has a *single* underlying stable growth trend (determined by its natural, human and capital resources, and the way they are utilised into production), and that shocks to earnings (or to output, employment, and/or population) are *temporary* deviations from this steady state. Under this view, resilience has to do with the ability to maintain stability near equilibrium during a recession.




*Source*: Martin (2012)
Impact of Recessionary Shocks on a Region's Growth Path:
Resumption to Pre-Shock Growth Trend

Therefore, from the 'engineering' perspective, a local economy is more resilient than another if it is less vulnerable when hit by the shock (e.g. earnings do not contract, or they do but with a relatively less pronounced impact) and rebounds more quickly after the shock to its *pre-existing* growth ceiling. Both resistance and recovery are thus important elements to economic resilience.

### *2.2.2 Equilibrium approaches to resilience: 'ecological' perspective*

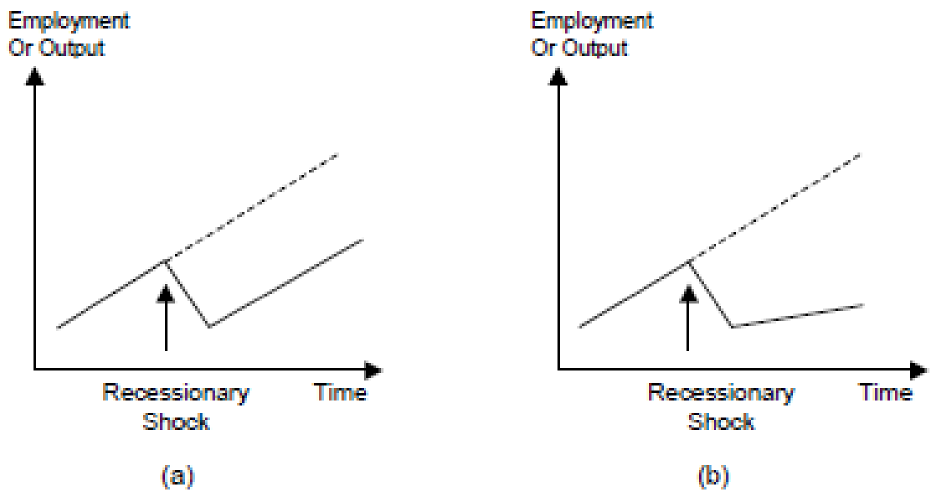
The 'ecological' version of economic resilience (Holling, 1973, 1996, 2001; McGlade et al., 2006; Walker et al., 2006) assumes that a local economy is characterised by *multiple* equilibria, and can transition from one to another as a result of a shock. In economics, this phenomenon is described using the notions of 'hysteresis' (Blanchard and Summers, 1987; Cross and Allen, 1988; Cross, 1993; Setterfield, 2010) or 'remanence' (Cross, Grinfeld and Lamba, 2009; Cross, Mcnamara and Pokrovskii, 2010). The idea here is that a deep or prolonged recession can cause a *permanent* downward or upward shift in the equilibrium growth trend, as opposed to the Plucking Model's assumption that effects are only temporary.

From the 'ecological' perspective, towns which resume growth but at a *new* lower level of, say, earnings (i.e. negative hysteretic effects) are deemed to be non-resilient or shock-prone; by contrast, towns which experience a full recovery, or move to a *new* superior steady state (i.e. positive hysteretic effects), are considered as economically resilient, or highly resilient.

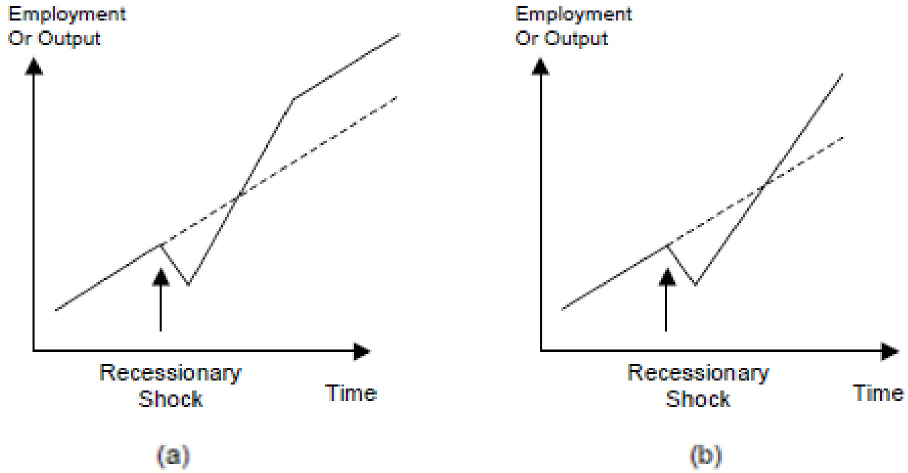
These neoclassical, equilibrium-based approaches are limited in their assumption of adjustment to a single or multiple stability domains. The 'engineering' interpretation of resilience as 'bounce-back' to an underlying stable growth trend ignores that an economy usually undergoes structural changes, whether following a shock or independently of it, and that these can influence the economy's resilience to future recessions. Resilience is thus an evolutionary process and not a static characteristic or feature of economies (Hill et al., 2008; Simmie and Martin, 2010; Martin, 2012). In this sense, however, the 'ecological' approach has an advantage over the 'engineering' approach since it recognises that major one-time disturbances can have effects which are left behind in the economy, and can permanently reduce or raise the economy's long-run growth path to an inferior or more desirable level. This implies that, as 'ecological' resilience is associated with the economic notions of hysteresis and remanence, there is in fact no need to relate it to the existence of equilibria (Setterfield, 2010).

Also, compared with the 'engineering' approach, the 'ecological' approach offers an opportunity to link resilience with the idea of adaptability, which is consistent with the view of resilience as an evolutionary process (McGlade et al., 2006). 'Engineering' resilience can be understood as the ability of an economy to *retain* its structures (industries, technologies, institutions, workforce) despite a shock rather than the ability to *change* them in response to the shock. By contrast, a resilient economy under the 'ecological' approach would be one that *adapts* successfully and either resumes or improves its long-run growth path; a non-resilient region would be one that fails to renew itself successfully and instead becomes 'locked' into outmoded structures, with a lowering of its long-run growth path. The relatively higher vulnerability to shocks of some towns can then be explained by their inability to adapt to the shock (and not to their inability to absorb it without any significant change, as in the 'engineering' view).

Nevertheless, both the 'engineering' and the 'ecological' perspectives consider resilience as recovery to a (pre-existing or new) steady state rather than as an ongoing process.




*Source*: Martin (2012)
Negative Hysteretic Impacts of Recessionary Shocks on a Region's Growth Path:
(a) Permanent Decline in Level, Resumption of Pre-Shock Growth Rate
(b) Permanent Decline in Level, and Lowered Growth Rate




*Source*: Martin (2012)
Hysteretic Impacts of Recessionary Shocks on a Region's Growth Path:
(a) Permanent Rise in Level, Resumption of Pre-Shock Growth Rate
(b) Permance Rise in Level, and Increased Growth Rate

***2.2.3 Evolutionary approaches to resilience:*** *concept of* ***'adaptability'***

Under the evolutionary-economics view (Martin and Sunley, 2006, 2009; Boschma and Martin, 2009; Simmie and Martin, 2010; Martin, 2012), resilience has to do with the adaptive capacity of a local economy to re-orientate institutions, resources, technologies and skills so as to preserve an acceptable growth path *over time*.

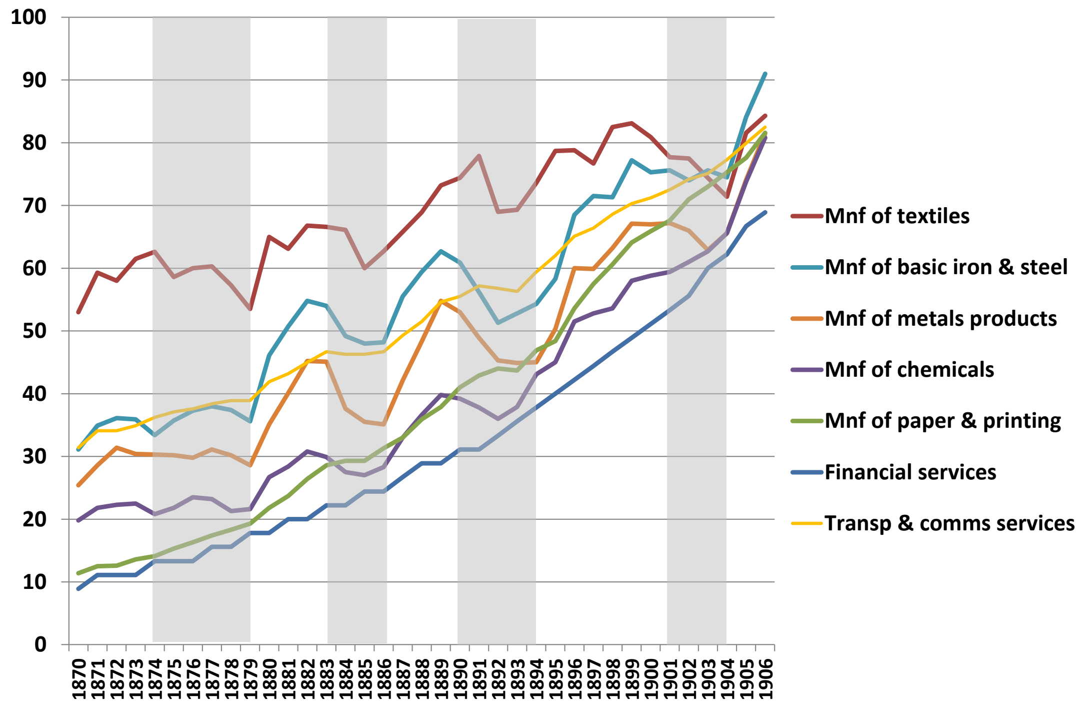
***2.2.4 Evolutionary approaches to resilience:*** *concept of* ***'lock-in'***

Adaptability is closely associated with the notions of path dependence and historical continuity which are at the core of the emergent evolutionary economic geography but go back to as early as David (1985) and Arthur (1989, 1990, 1994) in the form of lock-in analysis (see also David, 2005, 2007). The idea is that the location of an industry in a particular place can be determined by natural advantage (the presence of coal and minerals) or simply by initial 'historical accidents' or 'chance events'. Then, along the lines of Arthur's (1989) competing technologies model, what may happen is that various externality and learning mechanisms generate a process of 'cumulative causation' by which the early-established industry becomes ever more concentrated and new sectors cannot gain footing. Initially there is 'positive lock-in', a phase of industrial dynamism from close inter-firm relations. As the leading technology reaches maturity, however, the region enters a phase of 'negative lock-in' where it finds itself trapped into a development path which has become inferior or inefficient. The disruptive impact of an external shock such as a major downturn can destroy obsolete and unproductive activities and open up new growth opportunities, however those strong ties which previously underpinned local success become a source of inflexibility and lack of adaptability, ultimately leading to the town's economic decline.

The concepts of adaptability and lock-in are useful to the present discussion because, as well as explaining the evolution of the economic landscape and the process of regional development (i.e. differences in the long-run success of local economies), they also offer an explanation for why geographical areas differ in their vulnerability to and volatility during recessions.

### 2.3 Model and data

The modelling and forecasting exercises extend over the years before and after the 1890 shock which initiated the 1890-94 recession. The full sample period covers four slumps which can be dated (from peak to trough) to 1874-79, 1883-86, 1890-94 and 1901-04, taking the annual turning points of the UK business cycle as given in Aldcroft and Fearon (1972), Rostow (1948) and Southall (1986). The sectors being most affected by the 1890 shock (as well as earlier and subsequent shocks) were Britain's staple export industries, namely the manufacture of textiles, of iron and steel, and of metal products; all of the sample towns had some stake in these industries, which means that they were all potentially exposed to the negative effects of the shocks, although as it will be shown they did not react equally to them. By contrast, the contraction in output was much less severe for services and new growth industries such as printing which showed almost no signs of the downturns, as discussed in the sectoral analysis of historical data in Feinstein (1972).




*Source*: Feinstein (1972)
Real UK GDP by broad sector (1913=100)

### *2.3.1 Baseline specification*

This section shows the model of wage determination adopted to obtain parameter estimates for in-sample data (1872-1890). These are then used to give ex-post counterfactual wage predictions for the out-of-sample period (1890-1906) and to analyse spatial disparities in the impact of, and recovery from, the 1890 and subsequent shocks. Therefore, the ex-post counterfactual prediction exercise does not use out-of-sample data for the wage predictors but their time-autoregressive forecasts over the 1890-1906 period, constructed in such way so that they abstract from 1890 and earlier shocks and only reflect the underlying trend of the explanatory variables in question (see section 2.6). The counterfactual predictions of local wages are thus given by a linear combination of these forecasts with parameter estimates from the model illustrated in the present section, plus a correction term allowing for Best Linear Unbiased Predictions.

The *Great Britain Historical Database (GBHD)*[4] makes available thirty-five years of town-level[5] wage data, characterised by cyclical behaviour over time and uneven distribution in space. It thus gives $\text{Log}Wage_{i,t}$, which is the average nominal wage rate (pence per week[6]) of skilled engineering workers (i.e. 'fitters') in each of the sample towns (*i*) within Great Britain for each of the years (*t*) from 1872 to 1906. To summarise, the choice to use skilled engineering wages from the GBHD is driven by the availability of such spatial time series data, not readily obtainable from other sources. It also allows to operationalise the key concepts of the regional economic resilience literature, which are not confined to output growth but refer to any economic growth indicators such as population, employment or earnings (Martin, 2012); therefore, wages are used not as a proxy for output based on some theoretical link with output growth, but as one of many alternative measures of local economic performance which are explicitly envisioned by the reference literature.

The basis of the empirical model for **Log*Wage*** is provided by the extensive literature on the Wage Curve (Baltagi and Blien, 1998; Baltagi, Blien and Wolf, 2000, 2009; Bell, Nickell and Quintini, 2002; Blanchflower and Oswald, 1990, 1994a, 1994b, 1995, 2005; Buettner, 1999; Card, 1995; Fingleton and Longhi, 2013; Longhi et al., 2006; Nijkamp and Poot, 2005), which postulates an inverse relationship between the level of pay of individuals and the local unemployment rate (see review in Chapter 3). Given that historical time series of local unemployment rate are not available, I approximate the yearly unemployment for the sample towns by taking rates in 1868 (i.e. the year immediately before the start of the sample period) as reported in Southall (1986)[7] and by applying the annual change in UK

---

[4] The wage statistics are taken from Table SN3710 of the Great Britain Historical Database Online (Southall et al., 1999).

The original source of wage rates is an unpublished report on *Rates of Wages and Hours of Labour in various industries in the United Kingdom for a series of years* by the Board of Trade Labour Department (1908).

[5] Sample towns are Ashton-under-Lyne, Birmingham, Blackburn, Bolton, Bradford, Cardiff, Edinburgh, Greenock, Halifax, Hull, Leeds, London, Manchester, Newcastle, Nottingham, Sheffield, Sunderland, Wigan and Wolverhampton.

[6] There were 12 pence in a shilling, and 20 shillings or 240 pence in a pound.

[7] The original source of historical local unemployment rates is the Amalgamated Society of Engineers (A.S.E.) Monthly Reports available from the Mitchell Library (Glasgow), the British

unemployment[8]; this gives $\mathbf{Log}\boldsymbol{Unemp}$, the vector of town-specific series following the track of national unemployment. The model specification also includes two recession dummies, $\mathbf{Dummy}_{\mathbf{1874\text{-}79}}$ and $\mathbf{Dummy}_{\mathbf{1883\text{-}86}}$, to pick up the negative influence on local wages of the international banking crises which troughed in the late 1870s and in the mid-1880s. The additional covariate is the local mean wage, $\mathbf{Log}\boldsymbol{AvgWage}$, which is the average of wage rates across all sectors for which data are available; this variable is preferred to a linear time trend or national GDP as a business cycle measure as it captures cyclical output movements which are town-specific.

It should be noted that 'fitters' wages are appropriately excluded from this calculation, since their inclusion in the average would introduce two-way causation between the dependent variable and $\mathbf{Log}\boldsymbol{AvgWage}$. Nevertheless, some endogeneity concerns remain as this will possibly capture local spillovers from intra-town cross-sector linkages whereby productivity variations in one sector affect wage/productivity levels in other sectors within the same town, and thus it may be influenced by 'fitters' wages to some extent. The instrumentation strategy adopted in this essay is explained in section 2.4 with reference to both $\mathbf{Log}\boldsymbol{AvgWage}$ and $\mathbf{Log}\boldsymbol{Unemp}$.

### *2.3.2 Modelling spatial interaction*

The wage determination model proposed in the previous section is the reduced form of a more sophisticated and realistic specification, incorporating spatial effects, which is set out in the present and subsequent subsections. The baseline model rests on the assumption that wages in a typical town do not depend on wages in surrounding towns, and that a shock to the wage of, say, London has no impact on wages in other locations. However, in a system of open trading towns such as those in this sample, one should expect that spatial mechanisms have a role, and that local

---

Library of Political and Economic Science, the Trades Union Congress Library and Nuffield College (Oxford). For Wigan, which does not appear in the data, I take the unemployment rate of nearby Warrington.

[8] The source of historical national unemployment statistics is Feinstein (1972).

wages are to some extent determined by the characteristics of and developments in nearby towns as well as being the outcome of local labour-market conditions. One should also expect the existence of unobserved common factors driving wages in highly interrelated and interacting towns.

The estimation methodology in this essay draws on the burgeoning spatial panel modelling approach now becoming common in spatial econometrics. However, it is somewhat different in that it takes a flexible approach to what is meant by 'spatial' interdependence; that is, while in the literature inter-town proximity is typically based on some function of *geographical* distance, here it is also measured in a *socio-economic* sense following the suggestion in Corrado and Fingleton (2012) and Fingleton and Le Gallo (2008) of using similarity in town *population/income size* or local *employment structure*.

Meanwhile, the present study is driven by an appreciation that failure to acknowledge the presence of spatial effects would result in a misspecified model, and lead to an incorrect representation and understanding of the true causal processes at work. Model misspecification would have serious implications for the accuracy of econometric results and statistical inference (Le Sage and Pace, 2009). If the dependent variable exhibits spatial autocorrelation but the endogenous spatial lag is not included in the model, then coefficient estimates would be biased and inconsistent. Moreover, tests of hypotheses based on estimators which ignore spatial autocorrelation in the error term would give misleading outcomes; this is because leaving unobserved common factors (positive spatial residuals autocorrelation) unmodelled would lead to biased estimates of the variance of the regression parameters, reduced standard errors, inflated *t*-ratios and thus incorrect inference. Ultimately, neglecting cross-section dependence can cause counterfactual forecasts to be suboptimal or unreliable.

My approach to deal with spatial residuals dependence, by means of spatially autoregressive error components, allows to model network dependence explicitly. Since the impact of shocks is one focal point of the present essay, spatial effects operating through the error term should not be simply treated as nuisance (as in Spatial Heteroskedastic Autocorrelation Consistent estimation or common factor models); instead, I pay considerable attention to how shocks are transmitted across locations. Moreover, because I am interested in the spatial processes *per se*, I also seek to model spatial externalities or wage/productivity spillovers explicitly, and to achieve this a spatially autoregressive dependent variable is included which is obtained by pre-multiplying **Log*Wage*** by a non-stochastic spatial weights matrix; the presence of an endogenous spatial lag, here as in many other cases, is seen as necessary because it reflects true spatial linkages and "is not simply a surrogate for some omitted variables" (Corrado and Fingleton, 2012, p. 5).

Given this premise, and assuming spatial effects in the form of both an endogenous spatial lag and spatially autoregressive error components, I specify connectivity matrices $\mathbf{W}$ and $\mathbf{M}$ for the spatial lag and the error process respectively. These are $N$ x $N$ square matrices, with zeros on the main diagonal, which represent *a priori* hypotheses about the structure of connection between location pairs defined by specific rows and columns of $\mathbf{W}$ or $\mathbf{M}$; the (non-negative) value in any given cell $W_{ij}$ or $M_{ij}$ quantifies the hypothesised strength of interaction between towns *i* and *j*.

Below I give a more precise account of the treatment/modelling of spatial effects, while the next subsection presents the wage equation in detail.

### *2.3.3 Defining spatial weights matrices* $\mathbf{W}$ *and* $\mathbf{M}$

The use of different spatial weights matrices ($\mathbf{W}$ for the model's variables and $\mathbf{M}$ for the error process) is thoroughly motivated in a recent publication by Corrado and Fingleton (2012).

To construct $\mathbf{W}$, I adopt a <u>one nearest neighbour</u> weighting scheme, assuming that the level of wages in any town *i* is a positive function of the level of wages in its geographically closest town. Pace and Zou (2000) note that the nearest neighbour approach underfits the overall spatial dependence and provides a conservative estimate of the spatial structure; as it is apparent from Tables 2.6 and 2.7, the results shown later, the endogenous spatial lag based on this specification of the $\mathbf{W}$ matrix is highly significant so, even allowing fewer towns to interact, I find evidence of significant wage/productivity spillovers. Generally, this choice is appropriate given the relatively small cross-sectional dimension of the dataset at hand (only nineteen towns).

For **M** in Models 1 to 3 of Table 2.6 (section 2.5), I adopt a population-weighted inverse distance matrix. This is a well known and widely used definition of spatial connectivity in which the strength of dependence across towns inversely depends on the geographical distance between them and positively depends on the economic size (as measured by population or income) of the destination town. Distance and population values are rescaled, dividing them by a factor of one thousand, without loss of generality.

Moreover, all weights which are smaller than the global mean are set to zero (third statement below), thus assuming that the 'economic' separation between towns in each of the corresponding location pair is too large and so interactions are negligible. Using this cut-off, I avoid a full-distance matrix (i.e. a matrix with all weights being non-zero) which is necessary due to the asymptotics required to obtain consistent estimates for the parameters of the model[9]. In addition (fourth statement below), **M** is normalised by dividing each entry by the row sum, so that all rows add up to unity i.e. $\sum_{j=1}^{N} M_{ij} = 1$. Thus, the specification is as follows

$$
\begin{aligned}
M_{ij}^{*} &= \frac{POP_j}{d_{ij}} && i \neq j \\
M_{ij}^{*} &= 0 && i = j \\
M_{ij}^{*} &= 0 && M_{ij}^{*} < mean(M^{*}) \\
M_{ij} &= \frac{M_{ij}^{*}}{\sum_{j=1}^{N} M_{ij}^{*}}
\end{aligned}
$$

The resulting (non-symmetric, non-full) spatial weights matrix is reproduced in Table 2.1. One feature is that shocks are mostly transmitted from London and Manchester, which are the chief towns in the sample and thus disturbances to the path of wages in these towns are most likely to have repercussions elsewhere. Also, under the stated assumptions for **M**, Cardiff, Edinburgh and Greenock only receive *first-order* shocks from London. Moreover, although it is apparent that no shocks originate from these towns (and also from Hull and Nottingham) and that London is

[9] A full-distance matrix is usually not ideal because positive dependence for locations that are close in space averages out with negative dependence (e.g. based on some sort of hierarchical pattern) with locations further away. Thus, some cut-off has to be assumed, analogous to the case of the maximum lag length in temporal autocorrelation.

immune to shocks, all of these are indeed interested by *induced* contagion, as will become evident from the discussion about the Leontief expansion and spatial multiplier effects in the next subsection.

**Table 2.1 Normalised specification of M based on population-weighted inverse distance**

| | ASH | BIR | BLA | BOL | BRA | CAR | EDI | GRE | HAL | HUL | LEE | LON | MAN | NEW | NOT | SHE | SUN | WIG | WOL |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ASH | 0 | 0 | 0 | 0.079 | 0.100 | 0 | 0 | 0 | 0.069 | 0 | 0 | 0.464 | 0.228 | 0 | 0 | 0.061 | 0 | 0 | 0 |
| BIR | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.799 | 0.096 | 0 | 0 | 0 | 0 | 0 | 0.106 |
| BLA | 0 | 0 | 0 | 0.111 | 0.115 | 0 | 0 | 0 | 0.074 | 0 | 0 | 0.540 | 0.161 | 0 | 0 | 0 | 0 | 0 | 0 |
| BOL | 0.065 | 0 | 0.081 | 0 | 0.090 | 0 | 0 | 0 | 0.059 | 0 | 0 | 0.446 | 0.186 | 0 | 0 | 0 | 0 | 0.073 | 0 |
| BRA | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.148 | 0 | 0.138 | 0.575 | 0.139 | 0 | 0 | 0 | 0 | 0 | 0 |
| CAR | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| EDI | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| GRE | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| HAL | 0.057 | 0 | 0 | 0.059 | 0.188 | 0 | 0 | 0 | 0 | 0 | 0.082 | 0.437 | 0.119 | 0 | 0 | 0.057 | 0 | 0 | 0 |
| HUL | 0 | 0 | 0 | 0 | 0.109 | 0 | 0 | 0 | 0 | 0 | 0 | 0.782 | 0.109 | 0 | 0 | 0 | 0 | 0 | 0 |
| LEE | 0 | 0 | 0 | 0 | 0.198 | 0 | 0 | 0 | 0.093 | 0 | 0 | 0.526 | 0.114 | 0 | 0 | 0.069 | 0 | 0 | 0 |
| LON | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| MAN | 0.113 | 0 | 0.070 | 0.112 | 0.106 | 0 | 0 | 0 | 0.071 | 0 | 0 | 0.528 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| NEW | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.846 | 0 | 0 | 0 | 0 | 0.154 | 0 | 0 |
| NOT | 0 | 0.086 | 0 | 0 | 0.082 | 0 | 0 | 0 | 0 | 0 | 0 | 0.732 | 0.100 | 0 | 0 | 0 | 0 | 0 | 0 |
| SHE | 0 | 0 | 0 | 0 | 0.121 | 0 | 0 | 0 | 0 | 0 | 0.079 | 0.663 | 0.137 | 0 | 0 | 0 | 0 | 0 | 0 |
| SUN | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.735 | 0.112 | 0.153 | 0 | 0 | 0 | 0 | 0 |
| WIG | 0 | 0 | 0.085 | 0.123 | 0.094 | 0 | 0 | 0 | 0 | 0 | 0 | 0.526 | 0.172 | 0 | 0 | 0 | 0 | 0 | 0 |
| WOL | 0 | 0.173 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.728 | 0.099 | 0 | 0 | 0 | 0 | 0 | 0 |

In Models 4 and 6 of Table 2.7 (section 2.5), I use an alternative specification for **M** which is based on 'economic distance' (Corrado and Fingleton, 2012; Fingleton and Le Gallo, 2008). I make the hypothesis that flows are more intense between towns of similar size, thus assuming common shocks for locations that, although they may be geographically remote, are 'closer' in terms of size/population. This hypothesis is even more realistic for important economic centres, as in this sample, thanks to better communication infrastructure and lower transaction costs between them.

The variable used to compute pairwise one-dimension Euclidean distances is not absolute population but population rank (after sorting towns in ascending order by their population size at the start of the period, to ensure **M** is exogenous). Thus, the resulting measure of 'economic distance' is not the difference in the number of people living in any two towns $i$ and $j$ but the difference in their respective rank values (see Table 2.2).

The reason for this approach is that I am not interested in size *per se*, i.e. the difference in headcounts, but what I want to capture is the relative distance between towns in terms of how large or small they are. Therefore, for any town $i$, the weight carried by towns of relatively similar size will be higher, whereas the contribution from towns which are relatively far apart in terms of size ranking will be smaller as reflected by lower weights.

Not surprisingly, similarity in population size seems to reflect similarity in the sectoral composition of the local economic base through the industrial revolution period, with larger and diverse cities like London, Manchester or Birmingham at the top and smaller towns like the ports towns of Cardiff, Greenock, Hull and Sunderland at the bottom. Indeed, the link between town size and local economic structure is well documented in the urban economics literature, with larger towns found to be more diversified and smaller towns exhibiting higher specialisation (Duranton and Puga, 2000, 2001).

**Table 2.2 Town rank ($r$) based on population size**

| | |
|---|---|
| LONDON | 1 |
| MANCHESTER | 2 |
| BRADFORD | 3 |
| BIRMINGHAM | 4 |
| EDINBURGH | 5 |
| LEEDS | 6 |
| SHEFFIELD | 7 |
| BOLTON | 8 |
| HALIFAX | 9 |
| BLACKBURN | 10 |
| WOLVERHAMPTON | 11 |
| NEWCASTLE | 12 |
| ASHTON | 13 |
| NOTTINGHAM | 14 |
| SHUNDERLAND | 15 |
| WIGAN | 16 |
| CARDIFF | 17 |
| GREENOCK | 18 |
| HULL | 19 |

Table 2.3 illustrates the symmetric matrix of straight-line distances between the population size rank values of any two towns, where town ranks based on population size are as given in Table 2.2.

Hence, if $r_i$ and $r_j$ denote the rank positions of any two towns $i$ and $j$, the weight associated with this town pair will be the inverse of the one-dimension Euclidean distance between $r_i$ and $r_j$ i.e. $(Econ\_Dist_{ij})^{-1} = 1/\left|r_i - r_j\right|$, so that the definition of $\mathbf{M}$ adopted in Models 4 to 6 is

$$
\begin{array}{ll}
M_{ij}^{*} = (Econ\_Dist_{ij})^{-1} & i \neq j \\
M_{ij}^{*} = 0 & i = j \\
M_{ij}^{*} = 0 & M_{ij}^{*} < mean(M^{*}) \\
M_{ij} = \dfrac{M_{ij}^{*}}{\sum_{j=1}^{N} M_{ij}^{*}} &
\end{array}
$$

The resulting network dependence $\mathbf{M}$ matrix for the autoregressive error process (i.e. the inverse of that in Table 2.3) is illustrated in Table 2.4 (before normalisation) and in Table 2.5 (after normalisation, illustrated in full version); larger weights are assigned to pairs of towns – such as the largest, top-ranking towns of Manchester and London, or the smallest towns of Greenock and Cardiff or

Greenock and Hull - which are more similar in population size (and local economic structure).

**Table 2.3 Symmetric matrix of pairwise differences in rank positions based on 1871 population**

| | ASHT | BIRM | BLAC | BOLT | BRAD | CARD | EDIN | GREE | HALI | HULL | LEED | LOND | MANC | NEWC | NOTT | SHEF | SUND | WIGA | WOLV |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ASHT | 0 | 9 | 3 | 5 | 10 | 4 | 8 | 5 | 4 | 6 | 7 | 12 | 11 | 1 | 1 | 6 | 2 | 3 | 2 |
| BIRM | 9 | 0 | 6 | 4 | 1 | 13 | 1 | 14 | 5 | 15 | 2 | 3 | 2 | 8 | 10 | 3 | 11 | 12 | 7 |
| BLAC | 3 | 6 | 0 | 2 | 7 | 7 | 5 | 8 | 1 | 9 | 4 | 9 | 8 | 2 | 4 | 3 | 5 | 6 | 1 |
| BOLT | 5 | 4 | 2 | 0 | 5 | 9 | 3 | 10 | 1 | 11 | 2 | 7 | 6 | 4 | 6 | 1 | 7 | 8 | 3 |
| BRAD | 10 | 1 | 7 | 5 | 0 | 14 | 2 | 15 | 6 | 16 | 3 | 2 | 1 | 9 | 11 | 4 | 12 | 13 | 8 |
| CARD | 4 | 13 | 7 | 9 | 14 | 0 | 12 | 1 | 8 | 2 | 11 | 16 | 15 | 5 | 3 | 10 | 2 | 1 | 6 |
| EDIN | 8 | 1 | 5 | 3 | 2 | 12 | 0 | 13 | 4 | 14 | 1 | 4 | 3 | 7 | 9 | 2 | 10 | 11 | 6 |
| GREE | 5 | 14 | 8 | 10 | 15 | 1 | 13 | 0 | 9 | 1 | 12 | 17 | 16 | 6 | 4 | 11 | 3 | 2 | 7 |
| HALI | 4 | 5 | 1 | 1 | 6 | 8 | 4 | 9 | 0 | 10 | 3 | 8 | 7 | 3 | 5 | 2 | 6 | 7 | 2 |
| HULL | 6 | 15 | 9 | 11 | 16 | 2 | 14 | 1 | 10 | 0 | 13 | 18 | 17 | 7 | 5 | 12 | 4 | 3 | 8 |
| LEED | 7 | 2 | 4 | 2 | 3 | 11 | 1 | 12 | 3 | 13 | 0 | 5 | 4 | 6 | 8 | 1 | 9 | 10 | 5 |
| LOND | 12 | 3 | 9 | 7 | 2 | 16 | 4 | 17 | 8 | 18 | 5 | 0 | 1 | 11 | 13 | 6 | 14 | 15 | 10 |
| MANC | 11 | 2 | 8 | 6 | 1 | 15 | 3 | 16 | 7 | 17 | 4 | 1 | 0 | 10 | 12 | 5 | 13 | 14 | 9 |
| NEWC | 1 | 8 | 2 | 4 | 9 | 5 | 7 | 6 | 3 | 7 | 6 | 11 | 10 | 0 | 2 | 5 | 3 | 4 | 1 |
| NOTT | 1 | 10 | 4 | 6 | 11 | 3 | 9 | 4 | 5 | 5 | 8 | 13 | 12 | 2 | 0 | 7 | 1 | 2 | 3 |
| SHEF | 6 | 3 | 3 | 1 | 4 | 10 | 2 | 11 | 2 | 12 | 1 | 6 | 5 | 5 | 7 | 0 | 8 | 9 | 4 |
| SUND | 2 | 11 | 5 | 7 | 12 | 2 | 10 | 3 | 6 | 4 | 9 | 14 | 13 | 3 | 1 | 8 | 0 | 1 | 4 |
| WIGA | 3 | 12 | 6 | 8 | 13 | 1 | 11 | 2 | 7 | 3 | 10 | 15 | 14 | 4 | 2 | 9 | 1 | 0 | 5 |
| WOLV | 2 | 7 | 1 | 3 | 8 | 6 | 6 | 7 | 2 | 8 | 5 | 10 | 9 | 1 | 3 | 4 | 4 | 5 | 0 |

**Table 2.4 Non-normalised specification of M based on 'economic' distance**

| | ASHT | BIRM | BLAC | BOLT | BRAD | CARD | EDIN | GREE | HALI | HULL | LEED | LOND | MANC | NEWC | NOTT | SHEF | SUND | WIGA | WOLV |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ASHT | 0 | 0.111 | 0.333 | 0.2 | 0.1 | 0.25 | 0.125 | 0.2 | 0.25 | 0.167 | 0.143 | 0.083 | 0.091 | 1 | 1 | 0.167 | 0.5 | 0.333 | 0.5 |
| BIRM | 0.111 | 0 | 0.167 | 0.25 | 1 | 0.077 | 1 | 0.071 | 0.2 | 0.067 | 0.5 | 0.333 | 0.5 | 0.125 | 0.1 | 0.333 | 0.091 | 0.083 | 0.143 |
| BLAC | 0.333 | 0.167 | 0 | 0.5 | 0.143 | 0.143 | 0.2 | 0.125 | 1 | 0.111 | 0.25 | 0.111 | 0.125 | 0.5 | 0.25 | 0.333 | 0.2 | 0.167 | 1 |
| BOLT | 0.2 | 0.25 | 0.5 | 0 | 0.2 | 0.111 | 0.333 | 0.1 | 1 | 0.091 | 0.5 | 0.143 | 0.167 | 0.25 | 0.167 | 1 | 0.143 | 0.125 | 0.333 |
| BRAD | 0.1 | 1 | 0.143 | 0.2 | 0 | 0.071 | 0.5 | 0.067 | 0.167 | 0.063 | 0.333 | 0.5 | 1 | 0.111 | 0.091 | 0.25 | 0.083 | 0.077 | 0.125 |
| CARD | 0.25 | 0.077 | 0.143 | 0.111 | 0.071 | 0 | 0.083 | 1 | 0.125 | 0.5 | 0.091 | 0.063 | 0.067 | 0.2 | 0.333 | 0.1 | 0.5 | 1 | 0.167 |
| EDIN | 0.125 | 1 | 0.2 | 0.333 | 0.5 | 0.083 | 0 | 0.077 | 0.25 | 0.071 | 1 | 0.25 | 0.333 | 0.143 | 0.111 | 0.5 | 0.1 | 0.091 | 0.167 |
| GREE | 0.2 | 0.071 | 0.125 | 0.1 | 0.067 | 1 | 0.077 | 0 | 0.111 | 1 | 0.083 | 0.059 | 0.063 | 0.167 | 0.25 | 0.091 | 0.333 | 0.5 | 0.143 |
| HALI | 0.25 | 0.2 | 1 | 1 | 0.167 | 0.125 | 0.25 | 0.111 | 0 | 0.1 | 0.333 | 0.125 | 0.143 | 0.333 | 0.2 | 0.5 | 0.167 | 0.143 | 0.5 |
| HULL | 0.167 | 0.067 | 0.111 | 0.091 | 0.063 | 0.5 | 0.071 | 1 | 0.1 | 0 | 0.077 | 0.056 | 0.059 | 0.143 | 0.2 | 0.083 | 0.25 | 0.333 | 0.125 |
| LEED | 0.143 | 0.5 | 0.25 | 0.5 | 0.333 | 0.091 | 1 | 0.083 | 0.333 | 0.077 | 0 | 0.2 | 0.25 | 0.167 | 0.125 | 1 | 0.111 | 0.1 | 0.2 |
| LOND | 0.083 | 0.333 | 0.111 | 0.143 | 0.5 | 0.063 | 0.25 | 0.059 | 0.125 | 0.056 | 0.2 | 0 | 1 | 0.091 | 0.077 | 0.167 | 0.071 | 0.067 | 0.1 |
| MANC | 0.091 | 0.5 | 0.125 | 0.167 | 1 | 0.067 | 0.333 | 0.063 | 0.143 | 0.059 | 0.25 | 1 | 0 | 0.1 | 0.083 | 0.2 | 0.077 | 0.071 | 0.111 |
| NEWC | 1 | 0.125 | 0.5 | 0.25 | 0.111 | 0.2 | 0.143 | 0.167 | 0.333 | 0.143 | 0.167 | 0.091 | 0.1 | 0 | 0.5 | 0.2 | 0.333 | 0.25 | 1 |
| NOTT | 1 | 0.1 | 0.25 | 0.167 | 0.091 | 0.333 | 0.111 | 0.25 | 0.2 | 0.2 | 0.125 | 0.077 | 0.083 | 0.5 | 0 | 0.143 | 1 | 0.5 | 0.333 |
| SHEF | 0.167 | 0.333 | 0.333 | 1 | 0.25 | 0.1 | 0.5 | 0.091 | 0.5 | 0.083 | 1 | 0.167 | 0.2 | 0.2 | 0.143 | 0 | 0.125 | 0.111 | 0.25 |
| SUND | 0.5 | 0.091 | 0.2 | 0.143 | 0.083 | 0.5 | 0.1 | 0.333 | 0.167 | 0.25 | 0.111 | 0.071 | 0.077 | 0.333 | 1 | 0.125 | 0 | 1 | 0.25 |
| WIGA | 0.333 | 0.083 | 0.167 | 0.125 | 0.077 | 1 | 0.091 | 0.5 | 0.143 | 0.333 | 0.1 | 0.067 | 0.071 | 0.25 | 0.5 | 0.111 | 1 | 0 | 0.2 |
| WOLV | 0.5 | 0.143 | 1 | 0.333 | 0.125 | 0.167 | 0.167 | 0.143 | 0.5 | 0.125 | 0.2 | 0.1 | 0.111 | 1 | 0.333 | 0.25 | 0.25 | 0.2 | 0 |

**Table 2.5 Normalised specification of M based on 'economic' distance (full version)**

| | ASHT | BIRM | BLAC | BOLT | BRAD | CARD | EDIN | GREE | HALI | HULL | LEED | LOND | MANC | NEWC | NOTT | SHEF | SUND | WIGA | WOLV |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **ASHT** | 0 | 0.02 | 0.06 | 0.036 | 0.018 | 0.045 | 0.023 | 0.036 | 0.045 | 0.03 | 0.026 | 0.015 | 0.016 | 0.18 | 0.18 | 0.03 | 0.09 | 0.06 | 0.09 |
| **BIRM** | 0.022 | 0 | 0.032 | 0.049 | 0.194 | 0.015 | 0.194 | 0.014 | 0.039 | 0.013 | 0.097 | 0.065 | 0.097 | 0.024 | 0.019 | 0.065 | 0.018 | 0.016 | 0.028 |
| **BLAC** | 0.059 | 0.029 | 0 | 0.088 | 0.025 | 0.025 | 0.035 | 0.022 | 0.177 | 0.02 | 0.044 | 0.02 | 0.022 | 0.088 | 0.044 | 0.059 | 0.035 | 0.029 | 0.177 |
| **BOLT** | 0.036 | 0.045 | 0.089 | 0 | 0.036 | 0.02 | 0.059 | 0.018 | 0.178 | 0.016 | 0.089 | 0.025 | 0.03 | 0.045 | 0.03 | 0.178 | 0.025 | 0.022 | 0.059 |
| **BRAD** | 0.02 | 0.205 | 0.029 | 0.041 | 0 | 0.015 | 0.102 | 0.014 | 0.034 | 0.013 | 0.068 | 0.102 | 0.205 | 0.023 | 0.019 | 0.051 | 0.017 | 0.016 | 0.026 |
| **CARD** | 0.051 | 0.016 | 0.029 | 0.023 | 0.015 | 0 | 0.017 | ***0.205*** | 0.026 | 0.102 | 0.019 | 0.013 | 0.014 | 0.041 | 0.068 | 0.02 | 0.102 | 0.205 | 0.034 |
| **EDIN** | 0.023 | 0.187 | 0.037 | 0.062 | 0.094 | 0.016 | 0 | 0.014 | 0.047 | 0.013 | 0.187 | 0.047 | 0.062 | 0.027 | 0.021 | 0.094 | 0.019 | 0.017 | 0.031 |
| **GREE** | 0.045 | 0.016 | 0.028 | 0.023 | 0.015 | ***0.225*** | 0.017 | 0 | 0.025 | ***0.225*** | 0.019 | 0.013 | 0.014 | 0.038 | 0.056 | 0.02 | 0.075 | 0.113 | 0.032 |
| **HALI** | 0.044 | 0.035 | 0.177 | 0.177 | 0.03 | 0.022 | 0.044 | 0.02 | 0 | 0.018 | 0.059 | 0.022 | 0.025 | 0.059 | 0.035 | 0.089 | 0.03 | 0.025 | 0.089 |
| **HULL** | 0.048 | 0.019 | 0.032 | 0.026 | 0.018 | 0.143 | 0.02 | ***0.286*** | 0.029 | 0 | 0.022 | 0.016 | 0.017 | 0.041 | 0.057 | 0.024 | 0.072 | 0.095 | 0.036 |
| **LEED** | 0.026 | 0.092 | 0.046 | 0.092 | 0.061 | 0.017 | 0.183 | 0.015 | 0.061 | 0.014 | 0 | 0.037 | 0.046 | 0.031 | 0.023 | 0.183 | 0.02 | 0.018 | 0.037 |
| **LOND** | 0.024 | 0.095 | 0.032 | 0.041 | 0.143 | 0.018 | 0.072 | 0.017 | 0.036 | 0.016 | 0.057 | 0 | ***0.286*** | 0.026 | 0.022 | 0.048 | 0.02 | 0.019 | 0.029 |
| **MANC** | 0.02 | 0.113 | 0.028 | 0.038 | 0.225 | 0.015 | 0.075 | 0.014 | 0.032 | 0.013 | 0.056 | ***0.225*** | 0 | 0.023 | 0.019 | 0.045 | 0.017 | 0.016 | 0.025 |
| **NEWC** | 0.178 | 0.022 | 0.089 | 0.045 | 0.02 | 0.036 | 0.025 | 0.03 | 0.059 | 0.025 | 0.03 | 0.016 | 0.018 | 0 | 0.089 | 0.036 | 0.059 | 0.045 | 0.178 |
| **NOTT** | 0.183 | 0.018 | 0.046 | 0.031 | 0.017 | 0.061 | 0.02 | 0.046 | 0.037 | 0.037 | 0.023 | 0.014 | 0.015 | 0.092 | 0 | 0.026 | 0.183 | 0.092 | 0.061 |
| **SHEF** | 0.03 | 0.06 | 0.06 | 0.18 | 0.045 | 0.018 | 0.09 | 0.016 | 0.09 | 0.015 | 0.18 | 0.03 | 0.036 | 0.036 | 0.026 | 0 | 0.023 | 0.02 | 0.045 |
| **SUND** | 0.094 | 0.017 | 0.037 | 0.027 | 0.016 | 0.094 | 0.019 | 0.062 | 0.031 | 0.047 | 0.021 | 0.013 | 0.014 | 0.062 | 0.187 | 0.023 | 0 | 0.187 | 0.047 |
| **WIGA** | 0.065 | 0.016 | 0.032 | 0.024 | 0.015 | 0.194 | 0.018 | 0.097 | 0.028 | 0.065 | 0.019 | 0.013 | 0.014 | 0.049 | 0.097 | 0.022 | 0.194 | 0 | 0.039 |
| **WOLV** | 0.089 | 0.025 | 0.177 | 0.059 | 0.022 | 0.03 | 0.03 | 0.025 | 0.089 | 0.022 | 0.035 | 0.018 | 0.02 | 0.177 | 0.059 | 0.044 | 0.044 | 0.035 | 0 |

### *2.3.4 The SARAR-RE*[10] *Model specification*

This study draws from the spatial panel econometrics literature which started from Anselin (1988) and was recently pioneered by Kapoor, Kelejian and Prucha (2007), Fingleton (2008a, 2009a), Baltagi and Liu (2011), Baltagi (2013), Piras (2013) and Elhorst (2003a, 2010a, 2010b, 2012a, 2012b, 2013, 2014) (see also Baltagi, Song and Koh, 2003; Baltagi, Egger and Pfaffermayr, 2008).

The starting point is a panel data model with spatially and temporally autoregressive (as well as heteroskedastic) error components (with the autoregressive process also applying to the individual random effects) in combination with a spatially autoregressive dependent variable. This model is based on the work of Kapoor et al. (2007) but deals with spatial dependence also via an endogenous spatial lag rather than through the error process alone, as in Fingleton (2008a).

I begin with the panel specification

$$\mathbf{Y} = \mathbf{X}\boldsymbol{b} + \mathbf{e} \tag{2.1}$$

in which $\mathbf{Y}$ is an $NT$ x 1 vector of observations on (log) wages, given by $\mathbf{Log}\boldsymbol{Wage}_{it}$ for $i = 1...N$ and $t = 1...T$; $\mathbf{X}$ is an $NT$ x $k$ matrix of regressors comprising a constant, the (log) local unemployment rate, the (log) local average wage and the recession dummies; $\boldsymbol{b}$ is a $k$ x 1 vector of regression coefficients; $\mathbf{e}$ is an $NT$ x 1 vector given by a random error process. I introduce spatial effects both as a spatially autoregressive dependent variable and as a spatially autoregressive error process

$$\mathbf{Y} = \rho(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Y} + \mathbf{X}\boldsymbol{b} + \mathbf{e} = (\mathbf{I}_{TN} - \rho\mathbf{I}_T \otimes \mathbf{W})^{-1}(\mathbf{X}\boldsymbol{b} + \mathbf{e}) \tag{2.2}$$

$$\mathbf{e} = (\mathbf{I}_{TN} - \lambda\mathbf{I}_T \otimes \mathbf{M})^{-1}\boldsymbol{\xi} \tag{2.3}$$

where $\mathbf{I}_T$ is a $T$ x $T$ diagonal matrix with ones on the main diagonal and zeros elsewhere (identity matrix), $\mathbf{I}_N$ is a similar $N$ x $N$ diagonal matrix, $\mathbf{I}_{TN} = \mathbf{I}_T \otimes \mathbf{I}_N$ is an $NT$ x $NT$ diagonal matrix with ones on the main diagonal and zeros elsewhere, and $\boldsymbol{\xi}$ is an $NT$ x 1 vector of innovations. Also, $\rho$ and $\lambda$ are scalar parameters with $|\rho| < 1$ and $|\lambda| < 1$, $\mathbf{W}$ and $\mathbf{M}$ are $N$ x $N$ matrices of non-stochastic spatial weights

[10] Terminology often used in the literature, e.g. Anselin and Florax (1995).

which are row-normalised as in most spatial econometrics applications so that row sums equal one, and $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Y}$ is an $NT$ x 1 vector commonly referred to as an endogenous spatial lag.

To ensure stationarity and non-spurious regression, restrictions need to be imposed on $\mathbf{W}$ and $\mathbf{M}$ as well as on the scalar (spatial autoregressive) parameters $\rho$ and $\lambda$. One condition that should be satisfied before the spatial weight matrices are row-normalised is that the row and column sums of $\mathbf{W}$ and $\mathbf{M}$ are uniformly bounded in absolute value as $N$ goes to infinity, which means that a constant $c$ exists such that $\max_{1 \le i \le N} \sum_{j=1}^{N} \left|W_{ij}\right| \le c \le \infty$ and $\max_{1 \le j \le N} \sum_{i=1}^{N} \left|W_{ij}\right| \le c \le \infty$ (the same applies to $M_{ij}$ ) (Kelejian and Prucha, 1998; Kapoor et al., 2007). Another requirement is that $(\mathbf{I}_{TN} - \rho \mathbf{I}_T \otimes \mathbf{W})$ and $(\mathbf{I}_{TN} - \lambda \mathbf{I}_T \otimes \mathbf{M})$ are non-singular/ invertible; this condition is satisfied as long as $\rho$ and $\lambda$ are restricted to a stable continuous parameter space given by the inverse of the minimum and maximum characteristic roots of $\mathbf{W}$ or $\mathbf{M}$ (which need not be real values, see Kelejian and Prucha, 2010). After the matrices have been standardised by dividing each row cell by the row total, the largest i.e. most positive eigenvalue of $\mathbf{W}$ or $\mathbf{M}$ equals 1 while the smallest i.e. most negative eigenvalue can be less than -1 (Elhorst, 2010a). Thus $\rho$ should be in the interior of the following known parameter range (the feasible interval for $\lambda$ is analogue)

$$\frac{1}{\min(eig(\mathbf{W}))} < \rho < 1 = \frac{1}{\max(eig(\mathbf{W}))} \quad \text{for real characteristic roots, standardised } \mathbf{W} \quad (2.4)$$

The spatial autoregressive process for the error term, involving $(\mathbf{I}_{TN} - \lambda \mathbf{I}_T \otimes \mathbf{M})^{-1}$, implies complex instantaneous interdependence in wage levels across towns, so that a shock to the wage of town $i$ is simultaneously transmitted to all other towns and eventually works its way back to $i$. Taking just one cross-sectional regression at time $t$ and assuming $|\lambda| < 1$, the Leontief expansion shows that

$$\mathbf{e}_t = (\mathbf{I}_N - \lambda\mathbf{M})^{-1}\boldsymbol{\xi}_t = \left(\sum_{r=0}^{\infty}\lambda^r\mathbf{M}^r\right)\boldsymbol{\xi}_t = \\ = (\mathbf{I}_N + \lambda\mathbf{M} + \lambda^2\mathbf{M}^2 + \lambda^3\mathbf{M}^3 + \ldots)\boldsymbol{\xi}_t = \boldsymbol{\xi}_t + \lambda\mathbf{M}\boldsymbol{\xi}_t + \lambda^2\mathbf{M}^2\boldsymbol{\xi}_t + \lambda^3\mathbf{M}^3\boldsymbol{\xi}_t + \ldots \tag{2.5}$$

in which $\mathbf{M}^0 = \mathbf{I}_N$, $\mathbf{M}^2$ is the matrix product of $\mathbf{M}$ and $\mathbf{M}$, and $\mathbf{M}^r$ is the matrix product of $\mathbf{M}^{r-1}$ and $\mathbf{M}$. $\boldsymbol{\xi}_t$ is the *direct* effect that a shock to town *i* has on *i* while $\lambda\mathbf{M}\boldsymbol{\xi}_t$ is the *first-order indirect* effect that affects locations directly interacting with *i*, as given by the non-zero elements in $\mathbf{M}$. One can see the sum between direct and first-order indirect effects as local shock effects, and this is the type of shock transmission that occurs under a moving average (MA) error process. In contrast, the presence of the powers of $\mathbf{M}$ in the spatially autoregressive (SAR) error process implies that shocks are global; this is because a shock to town *i* affects *i*, the neighbours, the neighbours of the neighbours, and so on, cascading through all towns and eventually coming back to produce an additional (*induced*) effect on *i*. Therefore the full shock effect of the shock is the initial shock plus the feedback from all the other locations (Fingleton, 2008a). With the sample towns being open and highly-interconnected economies, it seems reasonable to assume that shocks are transmitted up and down the urban hierarchy and thus are global in nature. Moreover, $|\lambda| < 1$ implies diminishing importance of higher-order spatial lags (spatial linkages).

With some algebraic manipulation and re-arrangement, the specification in (2.2) can be expressed as

$$\mathbf{Y}_t = (\mathbf{I}_N - \rho\mathbf{W})^{-1}\mathbf{X}_t\boldsymbol{b} + (\mathbf{I}_N - \rho\mathbf{W})^{-1}(\mathbf{I}_N - \lambda\mathbf{M})^{-1}\boldsymbol{\xi}_t \tag{2.6}$$

which shows that the global spillover due to the SAR error process is amplified by the spatial multiplier effect due to the presence of the endogenous spatial lag. Anselin (2003, p. 11) notes that "the induced pattern of spatial dependence for the error term is much more complex and involves the interaction between the two spatial parameters as well as the two spatial weights".

For the space-time assumptions regarding the error components, I follow Kapoor et al. (2007), that is

$$\begin{aligned}&\mathbf{u} \sim iid(0, \sigma_u^2 \mathbf{I}_N)\\&\mathbf{v} \sim iid(0, \sigma_v^2 \mathbf{I}_N)\\&\boldsymbol{\xi} \sim (\boldsymbol{\iota}_T \otimes \mathbf{u}) + \mathbf{v}\end{aligned} \tag{2.7}$$

where $\mathbf{u}$ is an $N$ x 1 vector of errors specific to each area (i.e. the random effects), $\mathbf{v}$ is an $NT$ x 1 vector of errors specific to each area and each time with no covariance across area or time, $\boldsymbol{\iota}_T$ is a $T$ x 1 vector of ones, and $(\boldsymbol{\iota}_T \otimes \mathbf{u})$ is an $NT$ x 1 vector equal to $T$ stacked $\mathbf{u}'s$. Hence, for towns $i$, $j$ and times $t$, $s$:

$$\begin{aligned}&E(\xi_{it}\xi_{js}) = [\sigma_v^2 + \sigma_u^2] \quad && i = j \ ; \ t = s\\&E(\xi_{it}\xi_{js}) = [\sigma_u^2] && i = j \ ; \ t \neq s\\&E(\xi_{it}\xi_{js}) = [0] && i \neq j \ ; \ t \neq s\end{aligned} \tag{2.8}$$

with non-spherical innovations variance-covariance matrix $\boldsymbol{\Omega}_\xi$ (see Appendix 1). The model specification thus takes the form[11]

$$\begin{aligned}\mathbf{Log}\boldsymbol{Wage} = \boldsymbol{b}_0 + \rho(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{Wage} + \boldsymbol{b}_1\mathbf{Log}\boldsymbol{Unemp} +\\ \boldsymbol{b}_2\mathbf{Dummy}_{\mathbf{1874\text{-}79}} + \boldsymbol{b}_3\mathbf{Dummy}_{\mathbf{1883\text{-}86}} + \boldsymbol{b}_4\mathbf{Log}\boldsymbol{AvgWage} + \mathbf{e}\end{aligned} \tag{2.9}$$

$$\mathbf{e} = (\mathbf{I}_{TN} - \lambda\mathbf{I}_T \otimes \mathbf{M})^{-1}\boldsymbol{\xi} = [(\mathbf{I}_{TN} - \lambda\mathbf{I}_T \otimes \mathbf{M})^{-1}(\boldsymbol{\iota}_T \otimes \mathbf{u})] + [(\mathbf{I}_{TN} - \lambda\mathbf{I}_T \otimes \mathbf{M})^{-1}\mathbf{v}] \tag{2.10}$$

where $\mathbf{e}$ is an $NT$ x 1 vector of spatially dependent error terms, and $\boldsymbol{\xi}$ is an $NT$ x 1 vector of innovations which combines a permanent, i.e. time-constant, unit-specific error component $\mathbf{u}$ and a transient, i.e. time-varying, error component $\mathbf{v}$, respectively a random-effects vector picking up unobserved or unmeasured time-invariant interregional heterogeneity and a disturbances or shocks vector.

Therefore, the model's errors are both spatially and time-wise autocorrelated; time dependency is introduced into the innovations $\boldsymbol{\xi}$ by specifying the unobserved or unmeasured permanent unit-specific error component $\mathbf{u}$ together with the usual transient disturbance $\mathbf{v}$.

Importantly, this elaborated form of spatial dependence where the SAR process is not confined to $\mathbf{v}$ is standard in the cited spatial panel econometric

[11] $\mathbf{Log}\boldsymbol{Unemp}$ and $\mathbf{Log}\boldsymbol{AvgWage}$ are temporally lagged by one year, so that they pre-date the period of analysis and are pre-determined with respect to $\mathbf{Log}\boldsymbol{Wage}$, in order to address concerns about their potential endogeneity.

literature (e.g. Kapoor et al., 2007; Baltagi, Fingleton and Pirotte, 2014); its particular appeal is that each of the two error components **u** and **v** is subject to the 'same' spatially autoregressive process (see eq. 2.10), meaning that spatial autocorrelation in omitted (region-specific, permanent) explanatory variables is modelled explicitly.

To summarise, as necessary when analysing geographically-referenced data, the adopted model specification allows for:

- **'substantive' or 'systematic' spatial dependence** via a spatially lagged dependent variable representing spatial externalities or wage/productivity spillovers;
- **'nuisance' or 'non-systematic' spatial dependence** via (positive) spatial residuals autocorrelation due to transmission of global shocks across locations (common shocks) or (positive) spatial autocorrelation in unobserved or unmeasured causes of interregional heterogeneity (proximity effects, i.e. the fact that locations with similar socio-economic make-up are typically close to each other).

The specification accounts for town-specific **time-invariant** characteristics by means of random effects, as appropriate given the nature of the sample of towns at hand (which does not comprise all of the manufacturing towns in Great Britain but is chosen on the basis of data availability to represent that population). The random effects component, which is denoted by **u** and is part of the composite structure of the error term, is a catch-all for any causes of omitted (time-constant) spatial heterogeneity, including differences across towns in the sectoral composition of the local economic base and thus in their relative sensitivity to economic crises. By means of random effects the wage equation thus explicitly incorporates, as one of the possible determinants of local pay, variations in local industrial structure and in any other factors affecting the ability of towns to resist to and recover from recessions. This means that inter-town diversity in economic resilience to major shocks is built into the estimated model.

### 2.4 Instrumentation and estimation strategy

#### *2.4.1 Instrumental variables*

I have endeavored to exogenise $\mathbf{Log}\boldsymbol{AvgWage}$ and $\mathbf{Log}\boldsymbol{Unemp}$ by temporally lagging them and by excluding 'fitters' wages from the town average. On this basis, one can consider it reasonable to assume that $\mathbf{Log}\boldsymbol{AvgWage}$ and $\mathbf{Log}\boldsymbol{Unemp}$ are uncorrelated with the error term and thus with the dependent variable. Nevertheless, as previously stated, I also recognize that $\mathbf{Log}\boldsymbol{AvgWage}$ may be influenced by 'fitters' wages due to inter-sector spillovers within individual towns. In addition, one may have concerns about the quality of $\mathbf{Log}\boldsymbol{Unemp}$ itself, being based in part on modelled rather than observed data. Hence, to further ensure that estimates are not affected by endogeneity or measurement error, I obtain results from two estimators, first treating the variables as exogenous and then instrumenting them on the assumption that they are endogenous – and show that estimated parameters, and thus counterfactual predictions, are robust to whichever assumption I make.

- **Exogeneity assumption regarding RHS variables other than $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Y}$.**

In this case I follow the approach in Drukker, Egger and Prucha (2013) - who build on the econometric theory developed in Kelejian and Prucha (1998, 1999, 2004, 2010) – and use the lower orders of the spatial lags of the (included) exogenous variables $\mathbf{Log}\boldsymbol{Unemp}$ and $\mathbf{Log}\boldsymbol{AvgWage}$ to instrument the spatial lag variable (**Model 1** in Table 2.6 and **Model 4** in Table 2.7).

In empirical spatial econometric work, exogenous lags are widely accepted and well established as instrumental variables. In the above-cited literature, they are referred to as the recommended approach to instrumenting for an endogenous spatial lag, although the Monte Carlo simulations in Fingleton and Le Gallo (2007) have indicated that they are a good approximation of the optimal/ideal instruments also when applied outside the context of the spatial lag variable (this point is especially relevant in subsequent chapters where they are used to instrument variables other than the spatially lagged dependent variable).

As in Drukker, Egger and Prucha (2013), I define $\mathbf{X}_f = \left[\mathbf{X};\mathbf{X}_e\right]$ as the set of included exogenous variables and excluded exogenous variables, respectively $\mathbf{X}$ and

$\mathbf{X}_e$. Here $\mathbf{X}_f$ consists of $\mathbf{X} = \left[\mathbf{Log}\boldsymbol{Unemp}; \mathbf{Log}\boldsymbol{AvgWage}\right]$. The instruments are thus given by the linearly independent columns of

$$\mathbf{Z}=[\mathbf{X}_f; (\mathbf{I}_T \otimes \mathbf{W})\mathbf{X}_f; ...; (\mathbf{I}_T \otimes \mathbf{W})^q \mathbf{X}_f; (\mathbf{I}_T \otimes \mathbf{M})\mathbf{X}_f;\ (\mathbf{I}_T \otimes \mathbf{W})(\mathbf{I}_T \otimes \mathbf{M})\mathbf{X}_f;$$
$$...; (\mathbf{I}_T \otimes \mathbf{W})^q (\mathbf{I}_T \otimes \mathbf{M})\mathbf{X}_f]$$

With regard to the choice of *q*, a spatial lag order up to 2 has worked well in Monte Carlo simulations, but *q*=1 is a common choice in order to avoid an excessive proliferation of instrumental variables and severe multicollinearity, so I use the standard formulation of the instrument set with first-order spatial lags, that is

$$\mathbf{Z}=\left[\mathbf{X}_f; (\mathbf{I}_T \otimes \mathbf{W})\mathbf{X}_f; (\mathbf{I}_T \otimes \mathbf{M})\mathbf{X}_f; (\mathbf{I}_T \otimes \mathbf{W})(\mathbf{I}_T \otimes \mathbf{M})\mathbf{X}_f\right]$$
$$\mathbf{X}_f = \left[\mathbf{Log}\boldsymbol{Unemp}; \mathbf{Log}\boldsymbol{AvgWage}\right]$$

○ **Endogeneity assumption regarding RHS variables other than** $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Y}$.

Under the endogeneity assumption regarding $\mathbf{Log}\boldsymbol{Unemp}$ and $\mathbf{Log}\boldsymbol{AvgWage}$, their spatial lags are clearly endogenous (being a function of endogenous variables) and thus not valid instruments. In this case I adopt an IV set which is given by the 'three-group coding method'; this has a long and established track record having been used in quite a number of empirical applications, initially in the context of endogeneity induced by measurement error (Kennedy 2003) but also in a spatial framework (Fingleton 2003, Artis, Miguelez and Moreno, 2012).

The three-group method consists of coding each instrument with 1, 0 or -1 according to whether the values of the respective endogenous variable are in the top, middle or bottom third of their rankings. Here, the instruments set based on this approach thus includes the three-group transformation of $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{Wage}$, $\mathbf{Log}\boldsymbol{Unemp}$ and $\mathbf{Log}\boldsymbol{AvgWage}$ and the first-order spatial lags of $\text{3-Group}\,\mathbf{Log}\boldsymbol{Unemp}$ and $\text{3-Group}\,\mathbf{Log}\boldsymbol{AvgWage}$ (**Model 2** in Table 2.6 and **Model 5** in Table 2.7).

These 'quasi-instruments' are clearly relevant as they maintain correlation with the endogenous variables. However, strictly, the three-group variables constructed from endogenous regressors will keep an element of correlation with the errors and introduce some bias in estimation as pointed out by Fingleton and Le

Gallo (2007). However their Monte Carlo simulations indicate that they do not perform significantly worse than truly exogenous instruments and can thus be reliably used in applied work. Moreover it is reasonable to assume that the bias will be minimized by the combination of temporal and spatial lagging adopted here.

In **Model 3** (Table 2.6) and **Model 6** (Table 2.7), where the spatially lagged dependent variable is excluded, I use a variant of this IV set with no spatial lags of the three-group instrumental variables.

### *2.4.2 Fitting the SARAR-RE Model via S2SLS / GMM*

The models in this thesis are fitted by Spatial Two-Stage Least Squares / Generalised Method of Moments (S2SLS / GMM) instead of other methods such as Maximum Likelihood (ML) because ML cannot handle endogenous regressors other than the spatially lagged depedent variable. The S2SLS / GMM estimation procedure adopted throughout the thesis involves three stages.

1. In stage one, the model is estimated via Two-Stage Least Square (**2SLS**) to obtain consistent residuals, using instrument set $\mathbf{Z}$ as defined in the previous section.
2. In stage two, IV residuals from the first step are used in the non-linear Generalised Method of Moments (**GMM**) estimator of Kapoor et al. (2007) to obtain the model parameters relating to the error term, namely the error components variances $\sigma_\nu^2$ and $\sigma_u^2$ and the spatial-autoregressive error process parameter $\lambda$ (stage one uses arbitrary values of 1, 1 and 0 respectively for $\sigma_\nu^2$, $\sigma_1^2$ (where $\sigma_1^2 = \sigma_\nu^2 + T\sigma_u^2$) and $\lambda$).
3. Finally, given $\hat{\lambda}$, spatial dependence can be eliminated from the observed data and the error term by means of a **Cochrane–Orcutt transformation**, i.e. pre-multiplying by $(\mathbf{I}_{TN} - \hat{\lambda}(\mathbf{I}_{\boldsymbol{T}} \otimes \mathbf{M}))$. Using standard notation

$$\begin{aligned} \mathbf{Y}^* &= (\mathbf{I}_T \otimes (\mathbf{I}_N - \hat{\lambda}\mathbf{M}))\mathbf{Y} \\ \mathbf{X}_f^* &= (\mathbf{I}_T \otimes (\mathbf{I}_N - \hat{\lambda}\mathbf{M}))\mathbf{X}_f \\ \xi &= (\mathbf{I}_T \otimes (\mathbf{I}_N - \hat{\lambda}\mathbf{M}))\mathbf{e} \end{aligned}$$

Consistent estimates of structural parameters, namely the spatial autoregressive parameter $\rho$ and the other regression coefficients, can then be found based on **a robust approach for IV estimation with non-spherical disturbances**. Given innovations variance-covariance matrix $\Omega_{\xi} = E(\xi\xi') = \sigma_u^2(\mathbf{J}_T \otimes \mathbf{I}_N) + \sigma_\nu^2 \mathbf{I}_{TN}$ , which indicates that the disturbances are non-spherical and where $\mathbf{J}_T = \boldsymbol{\iota}_T \boldsymbol{\iota}_T'$ is a $T$ x $T$ matrix of ones, and given projection matrix $\mathbf{P}_Z = \mathbf{Z}(\mathbf{Z}'\hat{\Omega}_{\xi}\mathbf{Z})^{-1}\mathbf{Z}'$ where the instruments set $\mathbf{Z}$ is the same as in stage one, the vector of regression coefficients is

$$\begin{aligned}\hat{\boldsymbol{b}} &= \left[ (\mathbf{X}_f^{*\prime}\ \mathbf{Z}(\mathbf{Z}'\hat{\Omega}_{\xi}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}_f^{*}\ ) \right]^{-1} (\mathbf{X}_f^{*\prime}\ \mathbf{Z}(\mathbf{Z}'\hat{\Omega}_{\xi}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}^{*}) = \\ &= (\mathbf{X}_f^{*\prime}\ \mathbf{P}_Z \mathbf{X}_f^{*}\ )^{-1} \mathbf{X}_f^{*\prime}\ \mathbf{P}_Z \mathbf{Y}^{*}\end{aligned}$$

and the estimated variance-covariance matrix of regression coefficients is given by

$$\hat{\mathbf{C}} = \left[ (\mathbf{X}_f^{*\prime}\ \mathbf{Z})(\mathbf{Z}'\hat{\Omega}_{\xi}\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{X}_f^{*}) \right]^{-1} = (\mathbf{X}_f^{*\prime}\ \mathbf{P}_Z \mathbf{X}_f^{*}\ )^{-1}$$

The standard errors of the $\hat{\boldsymbol{b}}$ are given by the squares roots of the values on the main diagonal of $\hat{\mathbf{C}}$ , which allows $t$-ratios to be calculated for purposes of inference.

## 2.5 In-sample empirical results

### *2.5.1 Parameter estimates*

For Models 1 to 3 in Table 2.6 (where the structure of $\mathbf{M}$ is based on distance-decayed population size), it is apparent that the exogeneity/endogeneity assumption has no implications for the coefficient estimates and thus the counterfactual predictions; in contrast, Models 4 to 6 in Table 2.7 (where the spatial weights matrix $\mathbf{M}$ is based on 'economic distance') are less robust to the assumption about the exogeneity or endogeneity of $\mathbf{Log}\textbf{\textit{Unemp}}$ and $\mathbf{Log}\textbf{\textit{AvgWage}}$, as parameter estimates are somewhat worse under the exogeneity case. Under the endogeneity assumption, both specifications of $\mathbf{M}$ produce plausible results, with all coefficients statistically relevant and appropriately signed, although those from Model 2 are more in line with expectation across the board.

The model on which I choose to base the ex-post forecasting exercise is the one with the $\mathbf{M}$ structure which gives the 'best' (most robust, most plausible) set of results and the highest fit to the data, and which is the outcome of a more conservative estimation approach (treating regressors as potentially correlated with the error term and the dependent variable). The preferred model is thus Model 2, and in-sample results from this model are discussed in more detail in this section.

The estimate of the coefficient on $\mathbf{Log}\textbf{\textit{Unemp}}$ is appropriately signed and fairly significant (*t*-ratio=1.97) using typical inferential rules for two-tailed tests. In particular, a 1% rise (fall) in the local unemployment rate tends to reduce (increase) wages by approximately 0.03%. Therefore, although the unemployment elasticity of pay is lower than the 'empirical law' of (minus) 0.10%, a higher local joblessness rate is found to have a negative impact on local wages as suggested by the wage curve literature.

There are also strongly significant effects on wages of the general level of local wages (*t*-ratio=4.51). The estimated coefficient of 0.80 on $\mathbf{Log}\textbf{\textit{AvgWage}}$ implies that wages in a given town would increase by more than half if the general level of local wages doubled (i.e. $2^{0.7976}$-1=0.7382 or about 74%). This means that local wages are partly determined by cyclical movements in regional output/income, as one would expect, being driven up during expansions and pushed down during contractions of $\boldsymbol{AvgWage}$. This elasticity may also be the outcome of local spillovers from intra-town cross-sector linkages whereby productivity variations in one sector affect wage/productivity levels in other sectors within the same town.

The estimated coefficients on the two recession dummies are negatively signed, thus appropriately picking up the negative influence on local wages of the international banking crises which troughed in the late 1870s and in the mid-1880s; impact coefficients are respectively -0.0151 and -0.0259, implying that wages during the 1874-79 and 1883-86 slumps were respectively 1.5% and 2.6% lower than in normal times, with everything else constant. The two recession dummies have *t*-ratios of -2.18 and -2.87 respectively.

The relatively high value of the correlation between observed and fitted wages (0.8483) suggests that in-sample predictions are close to the original data and that the model fits the data well.

With regards to the two spatial effects in the model, there is evidence of positive interaction effects involving $\mathbf{Log}\boldsymbol{Wage}$, as suggested by $\hat{\rho}$ =0.3514 with a *t*-ratio of 1.93. The associated *P*-value for a two-tailed *t*-test with 355 degree of freedoms is 0.0545, which slightly exceeds the conventional 5% rate. However, assuming a one directional test, $\hat{\rho}$ has a *P*-value of 0.0273, which indicates a strong presence of the *positive* wage/productivity spillovers between any given town and its nearest neighbouring town, a hypothesis with a solid *a priori* rationale.

Therefore it is apparent that local wages are an outcome of between-town within-trade linkages, since they are a direct function of wages in the town in its closest proximity, and that high- (low-) wage towns tend to be clustered together in space. There is also evidence that the data generating process includes a positive autoregressive error process ($\hat{\lambda}$ =0.3668). The positive spatial correlation found in the error term is interpreted as the presence of global shock effects across connected towns, and these common shocks are instantaneously transmitted to all other locations with rebound effects on the towns they originated from.

To be noted is the fact that that $\hat{\lambda}$ is not significant in the absence of the spatial lag. On one hand, lower significance may be explained by the fact that the fit of the model is worse when the spatial lag is not a component of the equation (as indicated by the lower correlation between observed and predicted wages) and thus standard errors are increased. Another reason may be that, if the spatial lag should be in the model and is not, then one will have omitted variable bias in the parameter estimates, and this may be distorting $\hat{\lambda}$.

## Table 2.6 Results from M matrix based on distance-weighted population

| | 1 | 2 | 3 |
|---|---|---|---|
| W *matrix* | **Nearest Neighbour** | **Nearest Neighbour** | No Spatial Lag |
| M *matrix* | **Distance-weighted size/population** | **Distance-weighted size/population** | **Distance-weighted size/population** |
| *Treatment of* Log*Unemp* and Log*AvgWage* | **Exogenous** | **Endogenous** | **Endogenous** |
| *Instrument set* | **W*Log*Unemp*<br>W*Log*AvgWage*<br>M*Log*Unemp*<br>M*Log*AvgWage*<br>W*M*Log*Unemp*<br>W*M*Log*AvgWage*** | **3-Group Log*Unemp*<br>3-Group Log*AvgWage*<br>3-Group Wy<br>W*3-Group Log*Unemp*<br>W*3-Group Log*AvgWage*** | **3-Group Log*Unemp*<br>3-Group Log*AvgWage*** |
| Endog. Spatial Lag ($\rho$) [a]<br>(t-stat) | **0.4857**<br>(2.1958)** | **0.3514**<br>(1.9290)** | |
| Log*Unemp* ($b_1$)<br>(t-stat) | **-0.0223**<br>(-2.7054)*** | **-0.0275**<br>(-1.9658)** | **-0.0360**<br>(-2.7144)*** |
| Dummy 1874-79 ($b_2$)<br>(t-stat) | **-0.0145**<br>(-2.1763)** | **-0.0151**<br>(-2.1843)** | **-0.0157**<br>(-2.4319)** |
| Dummy 1883-86 ($b_3$)<br>(t-stat) | **-0.0252**<br>(-2.9967)*** | **-0.0259**<br>(-2.8723)*** | **-0.0242**<br>(-2.7832)*** |
| Log*AvgWage* ($b_4$)<br>(t-stat) | **0.7085**<br>(3.8955)*** | **0.7976**<br>(4.5128)*** | **0.9957**<br>(7.3802)*** |
| Constant<br>(t-stat) | -1.1315<br>(-1.2362) | -0.8610<br>(-0.9558) | 0.0580<br>(0.0738) |
| *Error process* | | | |
| $\lambda$ [b] | **0.3703*** | **0.3668*** | **0.2679** |
| $\sigma_v^2$ | 0.0009 | 0.0010 | 0.0010 |
| $\sigma_1^2 = \sigma_v^2 + T\sigma_u^2$ | 0.0485 | 0.0472 | 0.0444 |
| CORR [c] | **0.8443** | **0.8483** | **0.8209** |
| No. towns | 19 | 19 | 19 |
| No. in-sample years (1872-90) | 19 | 19 | 19 |
| No. observations | 361 | 361 | 361 |
| Degrees of freedom | 355 | 355 | 356 |

[a] Right-tailed test (null hypothesis: no *positive* effect of wages in the nearest neighboring town).
[b] Statistical significance of the estimated spatial autoregressive parameter is based on bootstrap inference.
[c] Correlation between observed and fitted values of **Log*Wage***.
* is significance at 10% level, ** is significance at 5% level, *** is significance at 1% level.

## Table 2.7 Results from M matrix based on 'economic' distance

| | 4 | 5 | 6 |
|---|---|---|---|
| *W matrix* | **Nearest Neighbour** | **Nearest Neighbour** | No Spatial Lag |
| *M matrix* | **'Economic' distance** | **'Economic' distance** | **'Economic' distance** |
| *Treatment of* Log*Unemp* and Log*AvgWage* | **Exogenous** | **Endogenous** | **Endogenous** |
| *Instrument set* | **W*Log*Unemp*<br>W*Log*AvgWage*<br>M*Log*Unemp*<br>M*Log*AvgWage*<br>W*M*Log*Unemp*<br>W*M*Log*AvgWage*** | **3-Group Log*Unemp*<br>3-Group Log*AvgWage*<br>3-Group Wy<br>W*3-Group Log*Unemp*<br>W*3-Group Log*AvgWage*** | **3-Group Log*Unemp*<br>3-Group Log*AvgWage*** |
| Endog. Spatial Lag ($\rho$) [a]<br>(t-stat) | **0.4965**<br>(2.6291)*** | **0.4099**<br>(2.7824)*** | |
| Log*Unemp* ($b_1$)<br>(t-stat) | **-0.0131**<br>(-1.8134)* | **-0.0302**<br>(-2.6900)*** | **-0.0358**<br>(-3.0763)*** |
| Dummy 1874-79 ($b_2$)<br>(t-stat) | **-0.0112**<br>(-1.5015) | **-0.0143**<br>(-1.7563)* | **-0.0157**<br>(-2.3940)** |
| Dummy 1883-86 ($b_3$)<br>(t-stat) | **-0.0109**<br>(-1.2206) | **-0.0169**<br>(-1.6922)* | **-0.0199**<br>(-2.2858)** |
| Log*AvgWage* ($b_4$)<br>(t-stat) | **0.6417**<br>(4.6150)*** | **0.7229**<br>(5.4384)*** | **0.9737**<br>(8.2704)*** |
| Constant<br>(t-stat) | -0.7991<br>(-1.0707) | -0.7343<br>(-1.0600) | 0.2019<br>(0.2935) |
| *Error process* | | | |
| $\lambda$ [b] | **0.4415***** | **0.4719***** | **0.2647** |
| $\sigma_v^2$ | 0.0010 | 0.0011 | 0.0011 |
| $\sigma_1^2 = \sigma_v^2 + T\sigma_u^2$ | 0.0261 | 0.0253 | 0.0337 |
| CORR [c] | **0.8411** | **0.8468** | **0.8209** |
| No. towns | 19 | 19 | 19 |
| No. in-sample years (1872-90) | 19 | 19 | 19 |
| No. observations | 361 | 361 | 361 |
| Degrees of freedom | 355 | 355 | 356 |

[a] Right-tailed test (null hypothesis: no *positive* effect of wages in the nearest neighboring town).
[b] Statistical significance of the estimated spatial autoregressive parameter is based on bootstrap inference.
[c] Correlation between observed and fitted values of **Log*Wage***.
* is significance at 10% level, ** is significance at 5% level, *** is significance at 1% level.

### *2.5.2 Bootstrap inference*

This section explains how a Bootstrap resampling scheme can be used to make inference about $\hat{\lambda}$, i.e. to generate an appropriate reference distribution for the spatial autoregressive parameter estimate and evaluate its significance.

The Bootstrap distribution is assumed to be close to the null distribution of $\hat{\lambda}$ when $\lambda = 0$ is true. Thus, the realised value of $\lambda$ is referred to its Bootstrap distribution to see how far out $\hat{\lambda}$ is on the right or left tail of the Bootstrap distribution with respect to the mean of this distribution.

More specifically, if $\hat{\lambda}$ is typical of values in the central portion of the Bootstrap distribution, then one would infer that the null hypothesis that $\lambda = 0$ should not be rejected. By contrast, if $\hat{\lambda}$ is an extreme occurrence under the null hypothesis that $\lambda = 0$, then one can consider the estimate to be significantly different from zero; for a directional test (null of no *positive* spatial residuals autocorrelation), if $\hat{\lambda}$ is sufficiently large that the proportion of times the Bootstrap replicates exceed $\hat{\lambda}$ is smaller than 5%, then one can infer that there is strong enough evidence against this null, with $\hat{\lambda}$ statistically significant at the 5% level.

The Bootstrap distribution is provided by resampling at random with replacement from IV residuals $\mathbf{e}_{IV}(\hat{\lambda})$. Thus, the $NT$ x 1 vector of IV residuals is the sampling frame, the sample size is equal to $NT$, and there is a probability equal to $1/NT$ of drawing the $i$th element of $\mathbf{e}_{IV}(\hat{\lambda})$.

**Figure 2.1 Distribution of $\hat{\lambda}$ from 'Equal Probability Systematic re-Sampling'**

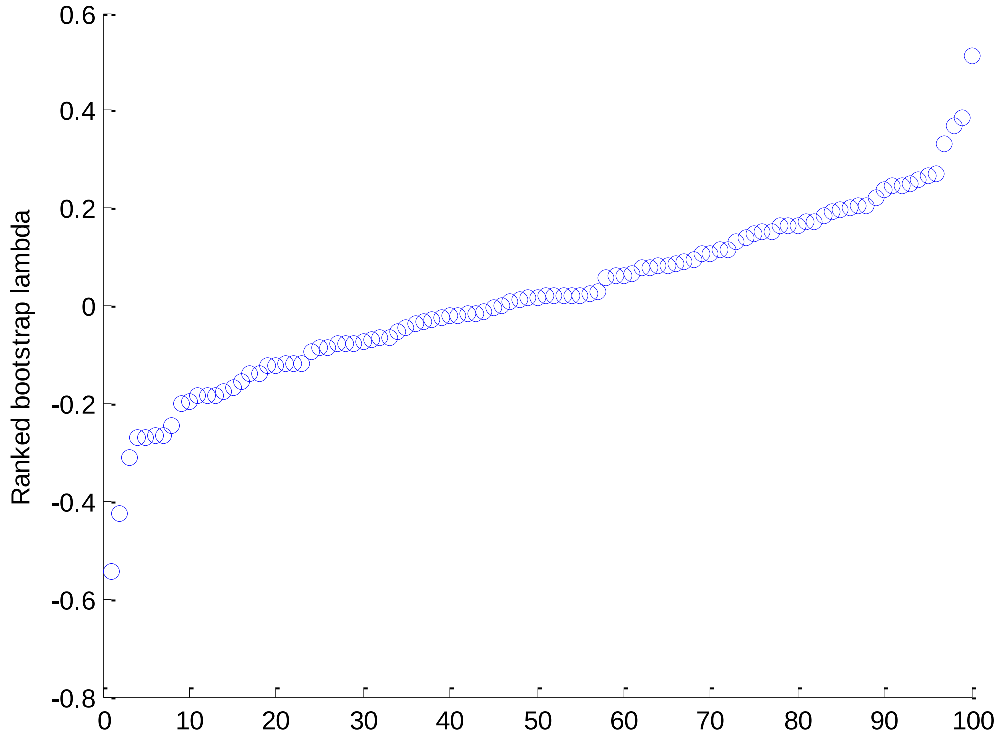



Figure 2.1 shows the empirical cumulative distribution function for $\hat{\lambda}$ based on one hundred iterations. Within this Bootstrap distribution, -0.23 and 0.23 are the critical values for a one-tailed probability of 8%, since eight out of one hundred observations lie to the left of -0.23 or to the right of 0.23; this means that any value of $\hat{\lambda}$ below -0.23 or above 0.23 is an unusual occurrence.

The estimation of Model 2 gives $\hat{\lambda}$=0.3668, which is above 0.23, ranking 98$^{th}$ among the sorted values of the Bootstrap distribution (from smallest to largest). Given that the Bootstrap distribution has a mean of 0.0184 and a standard deviation of 0.1764, the *t*-ratio associated with $\hat{\lambda}$=0.3668 is (0.3668-0.0184)/0.1764=1.98, meaning that $\hat{\lambda}$ is 1.98 standard errors away from its expected value under the null, which is zero. Thus, assuming a two-tailed Type I error rate of 5%, it is possible to reject the null in favour of a significant spatial autoregressive error process.

## 2.6 Out-of-sample (ex-post) counterfactual prediction exercise

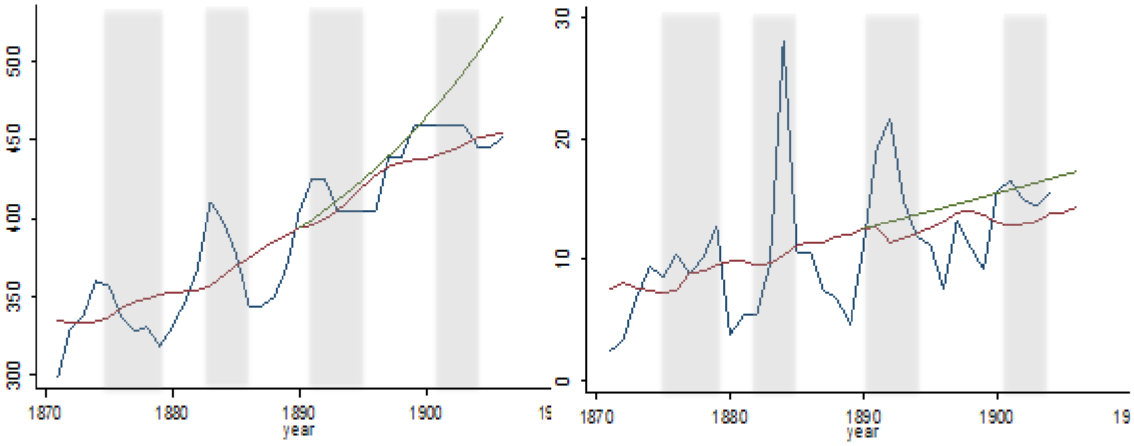
### *2.6.1 Forecasting methodology*

This section describes how counterfactual predictions of local wages are obtained.

First, for each individual town, I compute ex-post forecasts of local unemployment rate and local average wage over the out-of-sample period (1890-1906). These are generated via time-autoregressive models fitted to moving averages, rather than the actual levels of the explanatory variables, to better reflect the long-run trend in the time series for the two regressors. Thus, a seven-year moving average is applied to the local unemployment rate and average wage series prior to forecasting. The resulting time-autoregressive models provide out-of-sample forecasts which abstract from the economic cycle (i.e. the 1890 and earlier shocks); these are used in the prediction equation as reported below, thereby giving the estimated counterfactual wage levels.

I illustrate the forecasting methodology using data for Greenock, as one of the most volatile towns in the sample.

**Figure 2.2 Average wage (right) and unemployment rate (left) in Greenock**




*Notes*: (1) actual data: blue lines, seven-year moving averages: red lines, time-autoregressive forecasts (3 lags for **Log*AvgWage*** and 1 lag for **Log*Unemp***): green lines.

(2) **Log*AvgWage*** is pence per week, while **Log*Unemp*** is percentage of total population.

(3) Shaded areas denote UK recessions (1874-79, 1883-86, 1890-94, 1901-04), and highlight the negative correlation between wages and unemployment over the business cycle.

A seven-year moving average removes the peaks and troughs in the original data (blue lines in Figure 2.2), giving smoother series (red lines in Figure 2.2) which approximate the underlying trend (these series are not entirely smooth because moving averages leave short-term fluctuations and random noise intact). Time-autoregressive models are then fitted to these seven-year moving averages (i.e. actual data purged from shocks), choosing lag length on the basis of order-selection criteria, to obtain out-of-sample (counterfactual) forecasts (green lines in Figure 2.2).

Two points are worth mentioning with regard to this approach to trend estimation:

- Trend series are not directly observable and there is no universal agreement on the optimal technique to estimate them but any construction of these entails judgement. The choice of methods other than moving averages would require the application and confrontation of a number of alternative techniques, such as in the case of trend filtering, and the aim here is not to document estimates under different approaches to trend estimation.
- The choice of a moving-average window width of seven is appropriate given the sample size. Computationally, a seven-year moving average is equivalent to a local zero-degree polynomial smoother (unweighted) with bandwidth 0.45, i.e. to computing running means using 45% of the data. With 35 observations, this choice of bandwidth is at the lower bound of the common range of 0.4-0.8 for medium-sized samples (typical values for larger samples, i.e. $n$>=50, are 0.3-0.4).

For local unemployment rate, order selection criteria indicate an optimal lag length ($p$) of one for all towns; for local average wage, the lag length varies according to town, with results reported in Appendix 2.

I next combine the out-of-sample trend forecasts of explanatory variables with the regression coefficients from the estimated spatial panel data model (Model 2 in Table 2.6) to predict local wages under a no-recession scenario.

Finally the Baltagi, Bresson and Pirotte (2012) correction is added to obtain Best Linear Unbiased Predictions (BLUPs) for each individual town; more details are given in the next section.

### *2.6.2 Goldberger's (1962) BLUP correction*

Goldberger (1962) has shown that the Best Linear Unbiased Predictor (BLUP) for location *i* at a future period *T*+*s* is given by:

$$\hat{y}_{i,T+S} = x'_{i,T+s}\hat{b}_{S2SLS/GMM} + \boldsymbol{\omega}_i{}'\hat{\boldsymbol{\Omega}}_e^{-1}\hat{\mathbf{e}}_{S2SLS/GMM} \tag{2.11}$$

in which $\hat{y}_{i,T+S}$ is the scalar predicted value for *i* at *T*+*s*; $x'_{i,T+s}\hat{b}_{S2SLS}$ is the scalar expected value for *i* at *T*+*s* given the 1 x *k* vector of regressor values $x'_{i,T+s}$ and the *k* x 1 vector of coefficient estimates $\hat{\boldsymbol{b}}_{S2SLS}$; the scalar $\boldsymbol{\omega}_i{}'\hat{\boldsymbol{\Omega}}_e^{-1}\hat{\mathbf{e}}_{S2SLS/GMM}$ is the BLUP estimate of the prediction error for *i* at *T*+*s* with $\boldsymbol{\omega}_i{}'\hat{\boldsymbol{\Omega}}_e^{-1}$ being the 1 x *TN* correction term for location *i* (same for all years) - where $\boldsymbol{\omega}_i{}' = E\left[e_{i,T+s}e'\right]$ is the 1 x *NT* vector of covariances between the prediction error for *i* at *T*+*s* and the in-sample errors, $\hat{\boldsymbol{\Omega}}_e^{-1}$ is the *NT* x *NT* error variance-covariance matrix, and $\hat{\mathbf{e}}_{S2SLS/GMM}$ is the *NT* x 1 vector of in-sample residuals.

Baltagi, Bresson and Pirotte (2012) have demonstrated that the corrected prediction error for a spatial random-effects model *à la* Kapoor et al. (2007), i.e. with spatially autoregressive error components, as previously featured in Baltagi and Li (2006) and Fingleton (2009b) is

$$\boldsymbol{\omega}_i'\boldsymbol{\Omega}^{-1}\hat{\mathbf{e}} = \frac{\sigma_u^2}{T\sigma_u^2 + \sigma_v^2}(\boldsymbol{\iota}_T' \otimes l_i')\hat{\mathbf{e}} = \frac{\sigma_u^2}{\sigma_1^2}(\boldsymbol{\iota}_T' \otimes l_i')\hat{\mathbf{e}} \tag{2.12}$$

where $l_i$ is the *i*th column of $\mathbf{I}_N$ and $\boldsymbol{\iota}_T$ is a *T* x 1 vector of ones. I provide the full proof of eq. (2.12) in Appendix 1.

In eq. (2.12), the term $(\boldsymbol{\iota}_T' \otimes l_i')\hat{\mathbf{e}}$ is

$$\left[l_i'\ l_i'\ \dots\ l_i'\right]\hat{\mathbf{e}} = \left[l_i'\ l_i'\ \dots\ l_i'\right]\begin{bmatrix}\hat{e}_1\\ \hat{e}_2\\ \vdots\\ \hat{e}_T\end{bmatrix}$$

which equates to

$$\sum_{t=1}^{T} l_i'\hat{e}_t = \sum_{t=1}^{T}\hat{e}_{it} = T(\bar{\hat{e}})_i$$

where $(\bar{\hat{e}})_i$ is the average of $\hat{e}_{it}$ over t. Therefore, eq. (2.12) can be expressed as

$$\omega_i' \boldsymbol{\Omega}^{-1} \hat{\mathbf{e}} = \frac{\sigma_u^2}{T\sigma_u^2 + \sigma_v^2} (\boldsymbol{\iota}_T' \otimes l_i') \hat{\mathbf{e}} = T \frac{\sigma_u^2}{T\sigma_u^2 + \sigma_v^2} (\bar{\hat{e}})_i \quad (2.13)$$

which means that, by incorporating the Goldberger-type correction, I actually modify the usual SARAR-RE (in-sample and out-of-sample) prediction for town *i* by adding a fraction equal to $(T\,\sigma_u^2 / \sigma_1^2)$ of the mean of the residuals $\hat{\mathbf{e}}$ corresponding to the *i*th town (averaging over the *T* periods).

To obtain the out-of-sample counterfactual predictions of $\mathbf{Log}\boldsymbol{Wage}$, the parameter estimates $\hat{\rho}$ and $\hat{\boldsymbol{b}}$ from Model 2 of Table 2.6 are combined to the time-AR forecasts of $\mathbf{Log}\boldsymbol{Unemp}$ and $\mathbf{Log}\boldsymbol{AvgWage}$ (purged of recession effects by construction), with a BLUP correction applied to the error term.

However eq. (2.11) does not consider the fact that I actually have the spatial lag of the dependent variable in my estimation of eq. (2.9). Thus, to accommondate this, the Goldberger prediction equation becomes

$$\mathbf{Log}\boldsymbol{Wage}_{T+s} = \left(\mathbf{I}_N - \hat{\rho}\mathbf{W}\right)^{-1} \left[ \hat{\boldsymbol{b}}_0 \boldsymbol{\iota}_N + \hat{\boldsymbol{b}}_1 \mathbf{Log}\boldsymbol{Unemp}_{T+s} + \hat{\boldsymbol{b}}_4 \mathbf{Log}\boldsymbol{AvgWage}_{T+s} + \boldsymbol{\omega}' \hat{\boldsymbol{\Omega}}_e^{-1} \hat{\mathbf{e}} \right] \quad (2.14)$$

where *T*+*s* indicates any future time period and $\boldsymbol{\omega}' \hat{\boldsymbol{\Omega}}_e^{-1}$ is a 1 x *N* vector of corrections (same for all years).

Figures 2.3 and 2.4 plot observed wages against in-sample predictions with BLUP or non-BLUP errors, and show the improvement in the predictive performance of, respectively, SARAR-RE Model 2 and SARAR-RE Models 5 due to Goldberger’s correction.

In Figures 2.5 and 2.6, actual wage series for each individual town are plotted together with Goldberger-corrected fitted values (in-sample predictions) and counterfactual series (out-of-sample predictions) for, respectively, Model 2 and Model 5, to illustrate (after correction) how the estimated models were used to assess the impact of the shocks across towns.

**Figure 2.3 Fit of SARAR-RE Model 2 to in-sample data with and without Goldberger's (1962) BLUP correction**

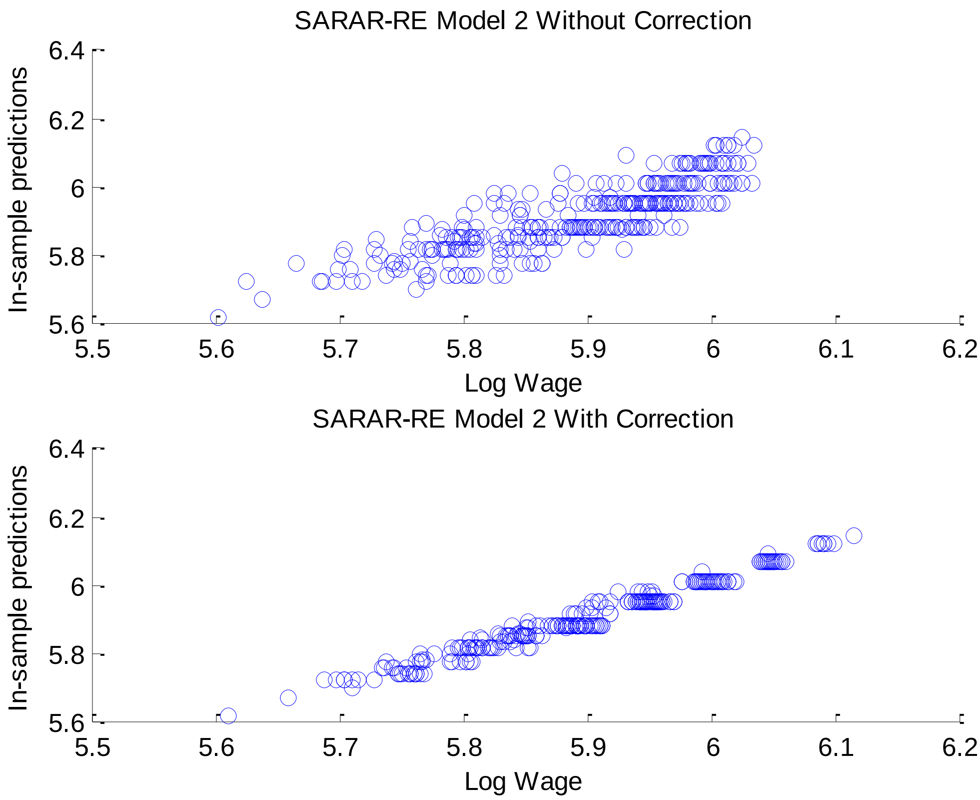


**Figure 2.4 Fit of SARAR-RE Model 5 to in-sample data with and without Goldberger's (1962) BLUP correction**

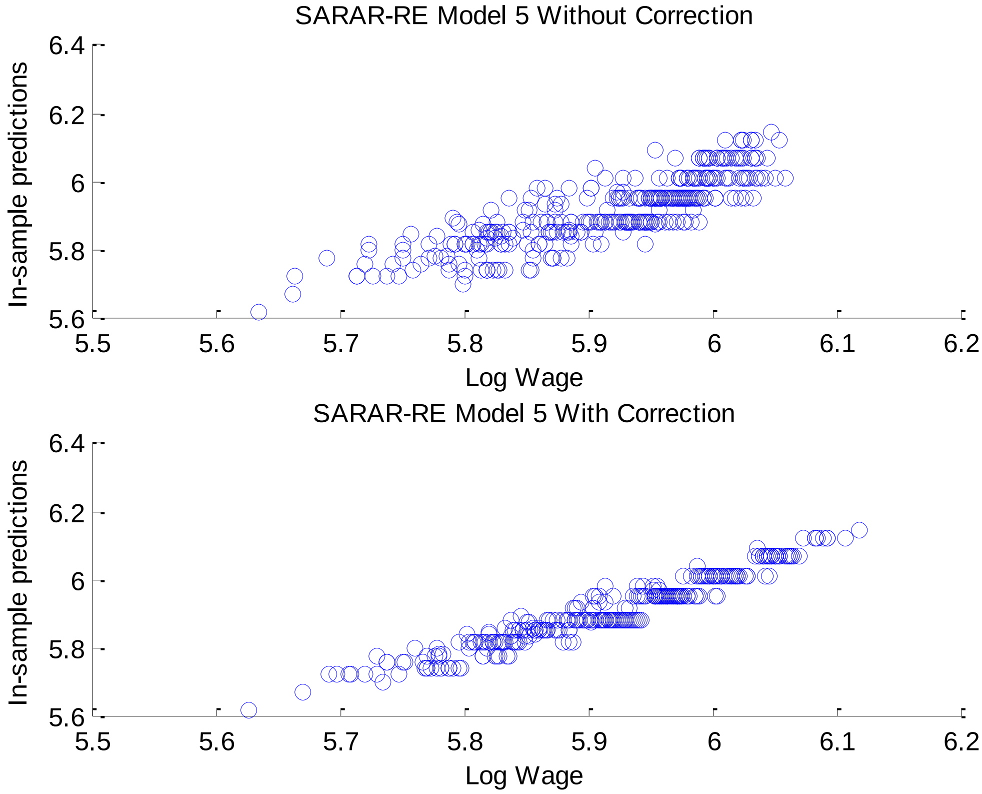
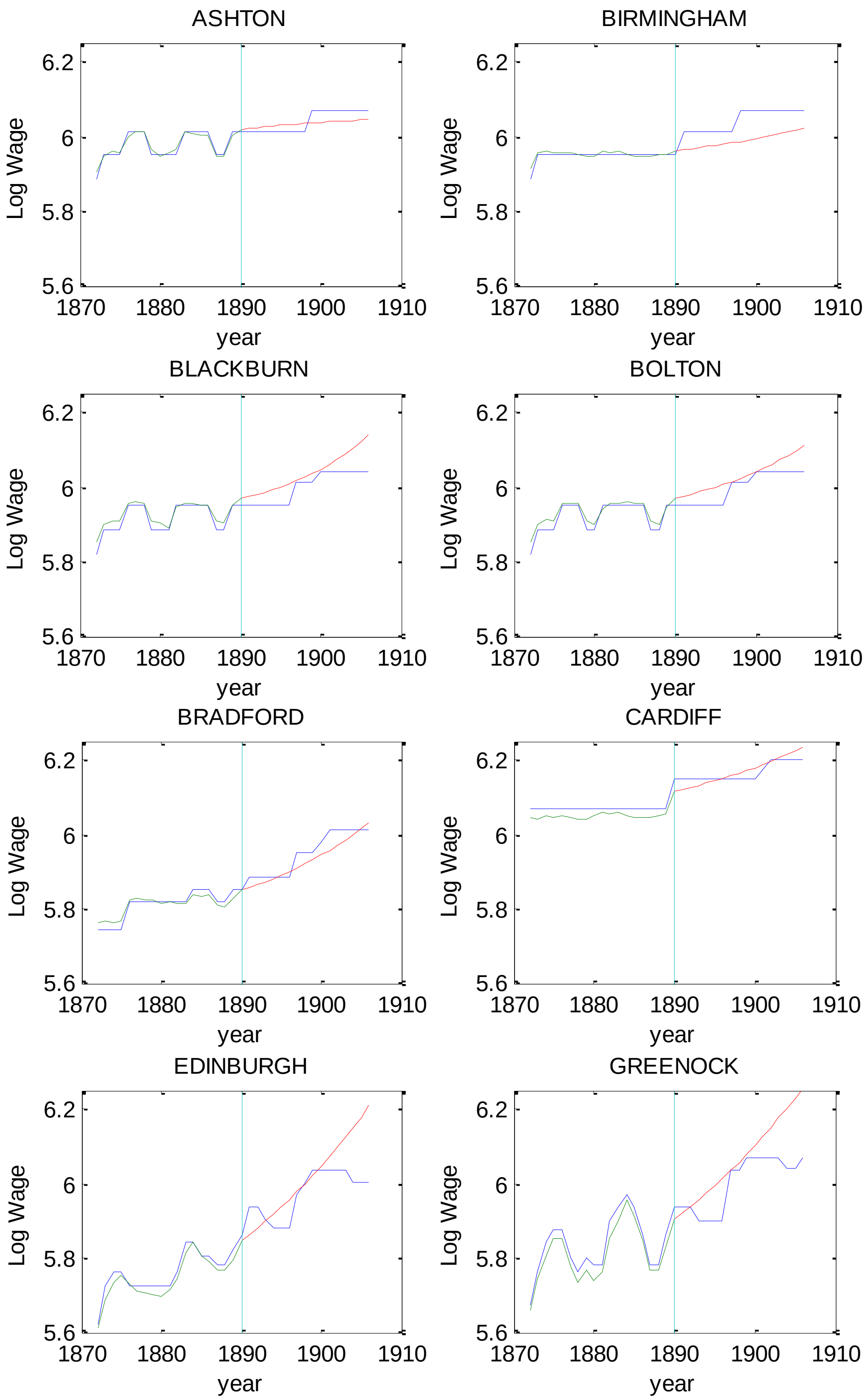
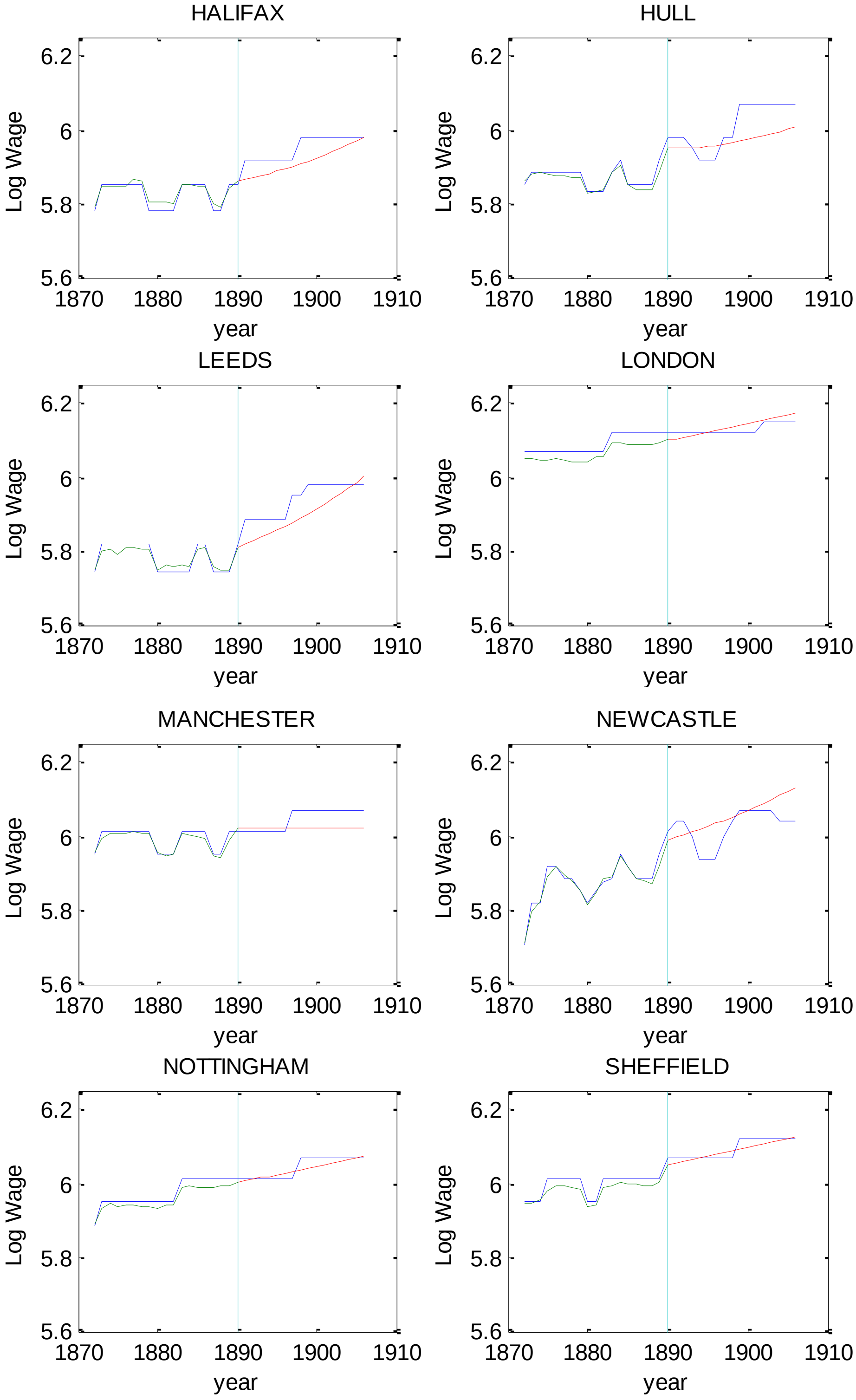


**Figure 2.5 Actual wages (blue line), and Goldberger-corrected in-sample predictions (green line) and out-of-sample predictions (red line) from SARAR-RE Model 2**




HALIFAX
HULL
LEEDS
LONDON
MANCHESTER
NEWCASTLE
NOTTINGHAM
SHEFFIELD
Log Wage
year
6.2
6
5.8
5.6
1870
1880
1890
1900
1910

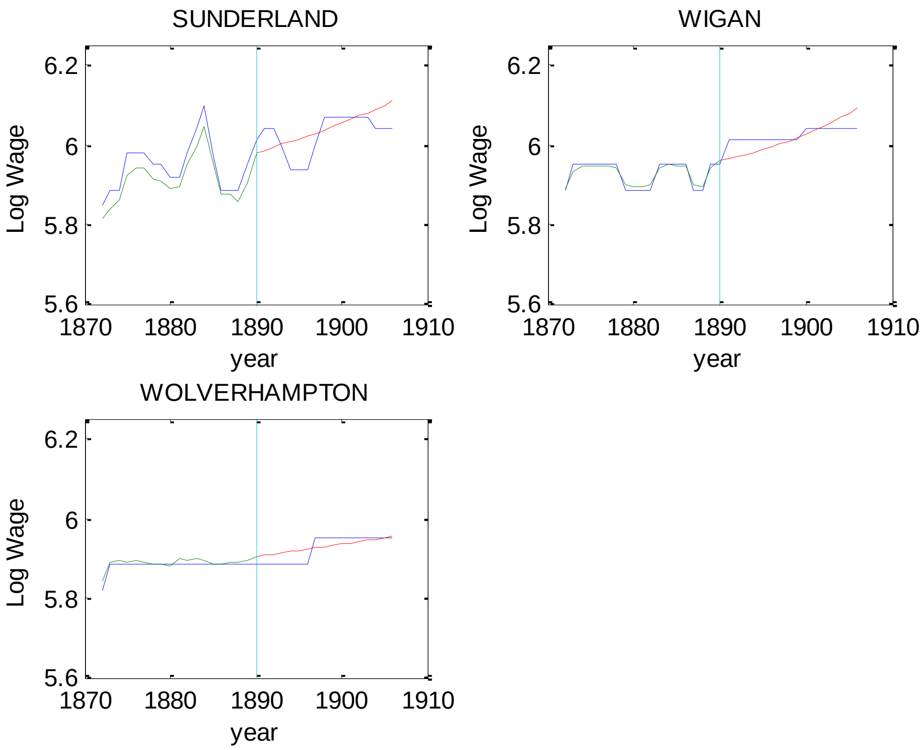
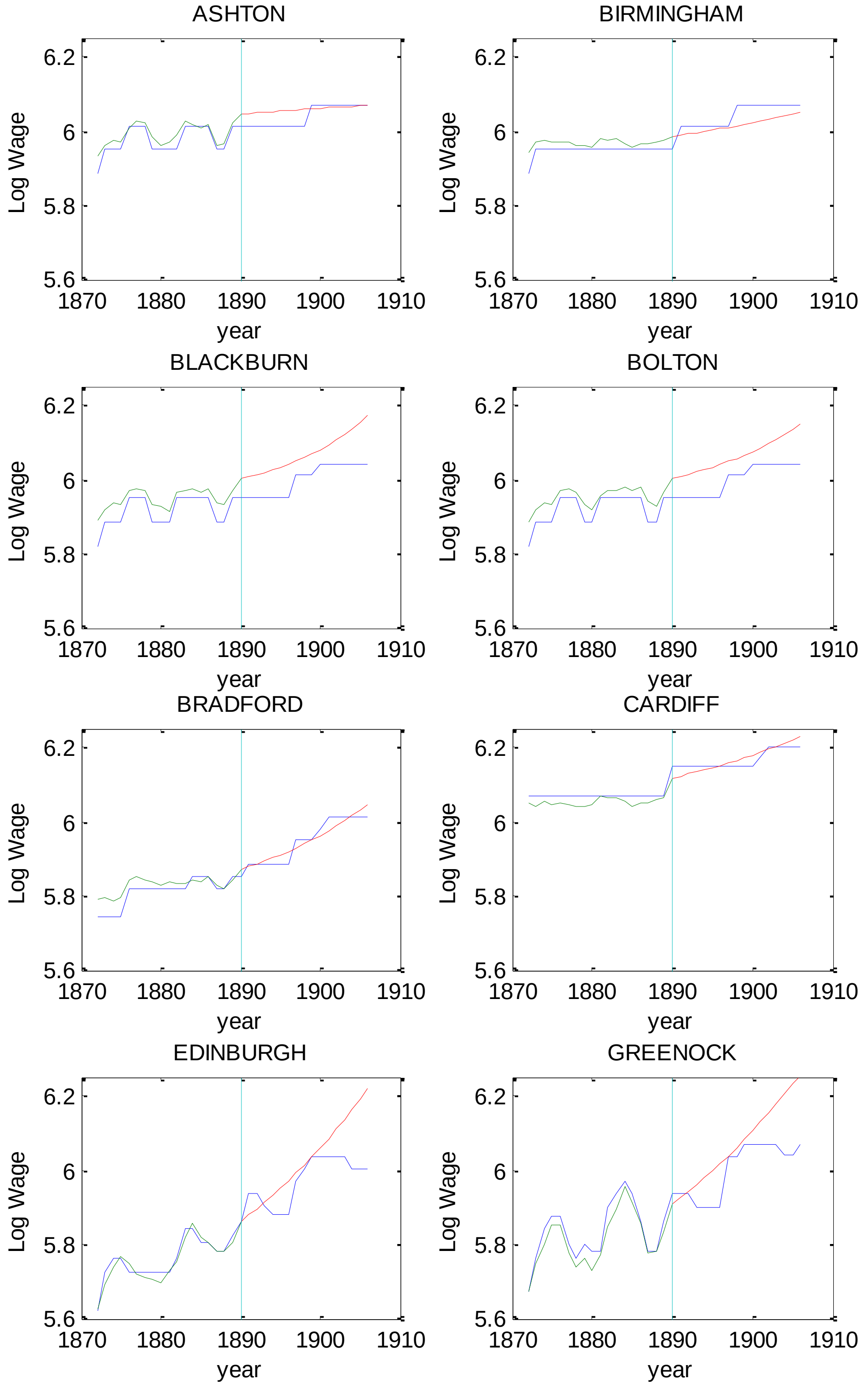
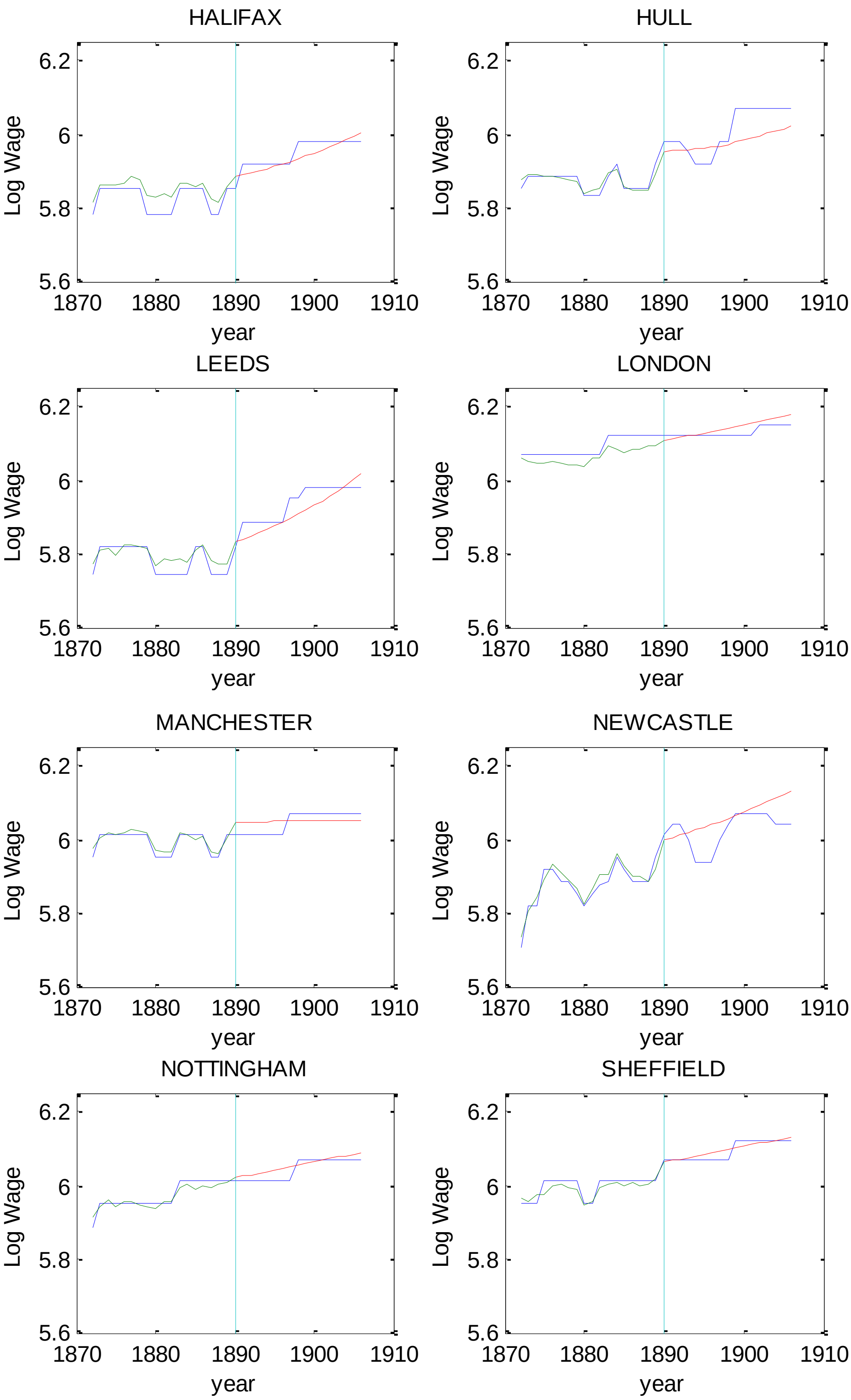
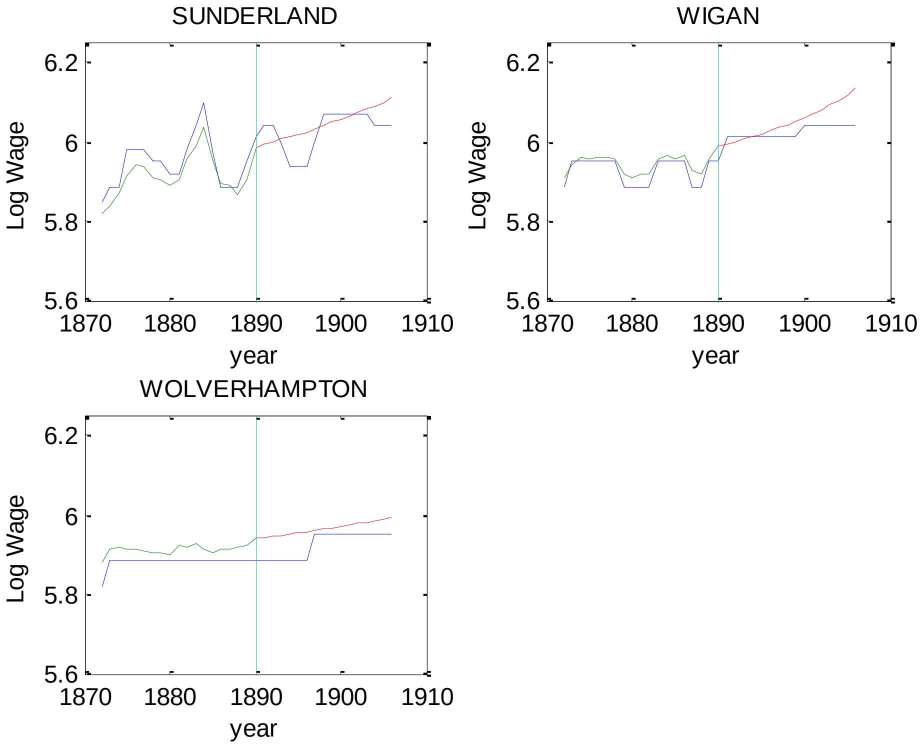

SUNDERLAND
WIGAN
WOLVERHAMPTON
Log Wage
year
6.2
6
5.8
5.6
1870
1880
1890
1900
1910

**Figure 2.6 Actual wages (blue line), and Goldberger-corrected in-sample predictions (green line) and out-of-sample predictions (red line) from SARAR-RE Model 5**




HALIFAX
HULL
LEEDS
LONDON
MANCHESTER
NEWCASTLE
NOTTINGHAM
SHEFFIELD
Log Wage
6.2
6
5.8
5.6
1870
1880
1890
1900
1910
year


SUNDERLAND
WIGAN
WOLVERHAMPTON
Log Wage
6.2
6
5.8
5.6
1870
1880
1890
1900
1910
year

## 2.7 Counterfactual analysis

### *2.7.1 Statistical framework for analysing resilience*

To gain insights into the economic resilience of the sample towns, this section presents results from the analysis of actual wages in relation to their counterfactual levels. The counterfactual wages are obtained using out-of-sample time-autoregressive trend forecasts for the regressors (unemployment rate and average wage) – which measure local economic conditions in a 'no-recession' scenario – and the coefficient estimates produced by SARAR-RE Model 2. The counterfactual predictions, which are BLUPs as they are corrected following Goldberger, represent the extension into the future of the pre-1890 trajectory of local wages.

The counterfactual analysis is based on two statistics. First (**Figure 2.7**) I measure, for each town in year 1906, by how much actual (log) wages were above or below counterfactual (log) wages; this gives an indication of whether, by the end of the 1890-1906 period, local wages had recovered to reach or exceed the counterfactual. A zero or very small difference implies regional labour market resilience in an 'ecological' sense, i.e. the ability of local wages to return ultimately to their long-run growth path; a relatively large positive difference suggests positive hysteretic effects, i.e. the ability of local wages to more than rebound, whereas a relatively large negative difference indicates negative hysteretic effects, i.e. a permanent downward shift in the equilibrium growth trend.

Second (**Figure 2.8**) I examine the entire out-of-sample period instead of focussing on an end-of-period snapshot. I calculate, for each town over the forecast years, the sum of the positive and negative deviations of actual (log) wages from counterfactual (log) wages; this measures whether wages observed after 1890 were on the whole above or below the level predicted for each town assuming that no shock had occurred. A zero or near-zero sum implies that local wages did not deviate, or not substantially deviate, from their counterfactual levels, therefore one would regard these towns as being resilient (in an 'engineering' sense); a relatively large negative value suggests that local wages were consistently below their counterfactual levels, whereas a relatively large positive value indicates a high degree of economic resilience.

**Figure 2.7 Difference between actual and counterfactual (log) wage levels in 1906**

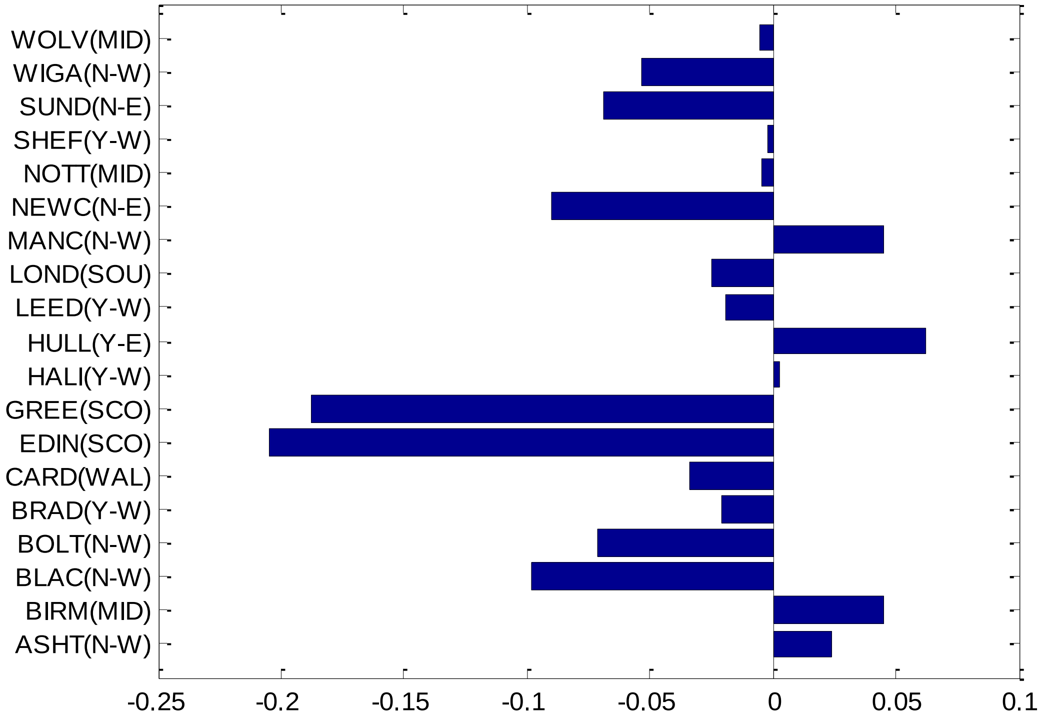
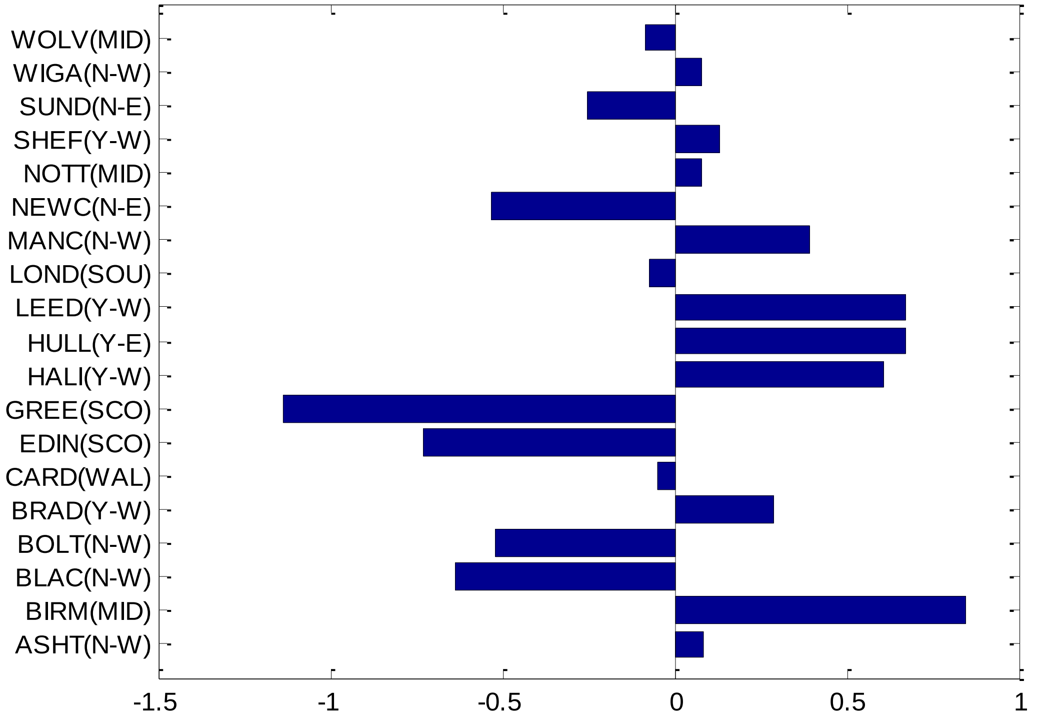



**Figure 2.8 Difference between actual and counterfactual (log) wage levels over the period 1890 to 1906**



### *2.7.2 Counterfactual analysis: results*

One finding from **Figure 2.7** is that, by 1906, local wages were in relative terms significantly below their counterfactual levels in Scotland (Edinburgh and Greenock) and in the North-East of England (Newcastle and Sunderland), all steel and shipbuilding centres apart from Edinburgh. The same is observed for the cotton towns of Lancashire (Blackburn, Bolton and Wigan), except for Manchester and nearby Ashton. These towns, where wages failed to return to their pre-1890 trajectory, thus show lack of economic resilience (i.e. negative hysteretic effects).

By contrast, in the chief seats for woollen and worsted manufactures of West Riding (Bradford, Halifax, Leeds and Sheffield), wages experienced relatively fast growth out of the recession, so that in 1906 they were close to their projected path. By 1906, wages had recovered or almost recovered also in Hull and Cardiff, two of the largest ports in Britain, in the Midlands (Birmingham, Nottingham and Wolverhampton) and in London. Therefore the interpretation is that wages in these towns were resilient or nearly resilient having recovered or almost recovered from the 1890 and subsequent shocks.

In **Figure 2.8** I see relatively large negative net differences in Scotland, the North-East and some towns in the cotton manufacturing heartland of Lancashire. The prominent double dip in local wages for Edinburgh, Greenock, Newcastle and Sunderland (Figure 2.6) indicates that the impact of the shock was relatively severe with initial recovery being choked off and growth turning negative again.

In contrast, I find that port towns outside Scotland and the North-Sea responded well, overall, to the negative effects of the 1890 and subsequent shocks. Likewise London and some towns in West Riding and in the Midlands, where on balance wages remained close to the counterfactuals, can be seen as resilient, with differences producing a comparatively small total over the 1890-1906 period.

The towns of Birmingham, Bradford, Halifax, Hull, Leeds and Manchester were highly resilient because, in relative terms, wages were on the whole significantly higher than expectation, with negative differences more than offset by positive differences, giving relatively large positive aggregate deviations from the counterfactuals. For example in Hull wages were initially relatively adversely impacted by the shock but recovered quickly (Figure 2.6).

Towns that had a positive balance relative to the counterfactual wage path, and which also had wages in 1906 above trend, one might consider to be 'super-resilient'. These are Manchester, Birmingham and Hull.

**Table 2.8 Town status (shock-prone, resilient, super-resilient)**

| | Recovery by 1906 (Figure 2.7 results) | Performance relative to the counterfactual over the period (Figure 2.8 results) | TOWN STATUS |
|---|---|---|---|
| WOLV | almost | similar | **RESILIENT** |
| WIGA | none | similar | **SHOCK-PRONE** |
| SUND | none | worse | **SHOCK-PRONE** |
| SHEF | almost | similar | **RESILIENT** |
| NOTT | almost | similar | **RESILIENT** |
| NEWC | none | worse | **SHOCK-PRONE** |
| **MANC** | **yes** | **better** | **SUPER-RESILIENT** |
| LOND | almost | similar | **RESILIENT** |
| LEED | almost | better | **RESILIENT** |
| **HULL** | **yes** | **better** | **SUPER-RESILIENT** |
| HALI | yes (but on trend, not above it) | better | **RESILIENT** |
| GREE | none | worse | **SHOCK-PRONE** |
| EDIN | none | worse | **SHOCK-PRONE** |
| CARD | almost | similar | **RESILIENT** |
| BRAD | almost | better | **RESILIENT** |
| BOLT | none | worse | **SHOCK-PRONE** |
| BLAC | none | worse | **SHOCK-PRONE** |
| **BIRM** | **yes** | **better** | **SUPER-RESILIENT** |
| ASHT | yes | similar (positive) | **SUPER-RESILIENT** |

Based on the statistics presented above and illustrated in Figures 2.7 and 2.8, Table 2.8 classifies sample towns as shock-prone, resilient and super-resilient. The highly specialised towns of Greenock, Newcastle, Sunderland, Blackburn, Bolton and Wigan are defined as *shock-prone*, whereas *resilient* places include the more diversified economies of West Riding, the Midlands and London. Among places classified as *super-resilient* are the cities of Birmingham and Manchester, both with extensive manufacturing activity and probably, as two of the largest towns in Britain, non-manufacturing trades also.

One interesting feature coming from this analysis is the lack of resilience of Blackburn, Bolton and Wigan compared to other textile centres in Lancashire, namely Manchester and nearby Ashton, together with Nottingham and West Riding towns. This possibly reflects the structural diversity of the latter economies (particularly Manchester and Nottingham) relative to the specialised, vulnerable towns elsewhere in the Lancashire cotton district, a topic I explore more fully below.

### *2.7.3 Counterfactual analysis: discussion*

Variety in the economic base has been suggested as a likely determinant of regional resilience by a number of authors (Hill et al., 2008; Martin, 2012). In the estimated model which produces the counterfactual predictions, I have captured town heterogeneity, including industrial structure, via random effects but these are time-constant effects, while in reality industrial structure in particular is a dynamic entity. This is clearly illustrated by the county-level data in Table 2.9, which show how some counties and probably the towns they embodied were becoming increasingly, or less, specialised over the period 1871 to 1911. For example Newcastle, which at the time was included in Northumberland, and Greenock [12], formerly part of Strathclyde, both were major and increasingly specialised shipbuilding towns while Warwickshire, which historically included much of Birmingham, was increasingly specialised in vehicle manufacture. So while the counterfactuals capture, via the random-effects error component, the 'average' ***static effect*** of industrial structure during the 1871-1890 period, they do not incorporate the ***dynamic changes*** that were taking place in the sectoral composition of local employment. Therefore the differential impact of the 1890 and subsequent shocks across the sample towns (as inferred from the performance of local wages with respect to their counterfactual levels) can be explained by their ***increasing or diminishing***, as well as ***excessive or lacking***, specialisation in specific industries. Most notably, increasing specialization seems to be associated with heightened exposure to shock effects.

For the Midlands, Table 2.9 shows that the region had low *or* falling specialisation in the old, crisis-hit metal and textile industries but had also successfully established new growth sectors such as vehicles (cars and cycles) (Hunt, 1973). Locally available skills and technologies are likely to have played a role also

[12] The town's shipyards included famous names such as Browns, Scotts and William Lithgows, supported by highly specialised marine engineering and ship repairing companies.

(Fingleton, Garretsen and Martin, 2012; Martin, 2012), because the region’s electrical and mechanical manufacturing developed from existing engineering activities such as the making of railway carriages and wagons and the manufacture of lace and textile machinery. All of this implied relatively high adaptability of the regional economy as redundant engineers in mature, staple industries could be relatively easily re-employed in the remaining, more favoured sectors. Birmingham emerges as super-resilient, thanks to its exceptionally diversified industrial mix and strong business and innovation culture as documented in Hunt (1973).

**Table 2.9 Location Quotients for selected Counties (employment share in each sector over national share)**

| Region, county | town | Metals 1871 | Metals 1911 | Printing 1871 | Printing 1911 | Services 1871 | Services 1911 | Shipbuilding 1871 | Shipbuilding 1911 | Textiles 1871 | Textiles 1911 | Vehicles 1871 | Vehicles 1911 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SCOTLAND | | | | | | | | | | | | | |
| Strathclyde | Greenock | 1.70 | 2.14 | 1.16 | 0.99 | 0.61 | 0.67 | *4.23* | *5.47* | 1.60 | 0.97 | 0.43 | 0.66 |
| Lothian | Edinburgh | 0.90 | 0.68 | *3.87* | *2.64* | *1.19* | *1.07* | 0.42 | 0.56 | 0.21 | 0.31 | 0.70 | 0.34 |
| LANCASHIRE | | | | | | | | | | | | | |
| Lancashire | Wigan,Manc's, Bolton,B'burn, Ashton | 0.90 | 0.87 | 0.89 | 0.93 | 0.74 | 0.68 | 0.86 | 0.50 | *2.89* | *3.16* | 0.87 | 0.67 |
| YORKSHIRE | | | | | | | | | | | | | |
| Yorkshire (West Riding) | Sheff'd,Bradf'd, Leeds,H'fax | 1.29 | 1.49 | 0.60 | 0.71 | 0.63 | 0.65 | 0.11 | 0.09 | *2.76* | *2.57* | 0.83 | 0.74 |
| Yorkshire (East Riding) | Hull | 0.89 | 0.77 | 0.73 | 0.82 | 1.16 | 0.99 | *2.39* | *1.80* | 0.19 | 0.18 | 1.56 | 1.03 |
| NORTH-EAST | | | | | | | | | | | | | |
| Durham | Sunderland | 3.04 | 1.69 | 0.57 | 0.44 | 0.68 | 0.68 | *5.79* | *6.57* | 0.13 | 0.17 | 0.51 | 0.44 |
| Northumberland | Newcastle | 1.30 | 0.96 | 0.78 | 0.48 | 0.95 | 1.90 | *3.88* | *5.61* | 0.14 | 0.16 | 0.65 | 0.43 |
| MIDLANDS | | | | | | | | | | | | | |
| Nottinghamshire | Nottingham | *0.62* | *0.71* | 0.73 | 0.97 | 0.78 | 0.75 | 0.04 | 0.03 | *2.24* | *2.04* | *1.10* | *1.03* |
| Staffordshire | Wolverh'pton | *3.10* | *2.55* | 0.30 | 0.51 | 0.76 | 0.69 | 0.19 | 0.06 | *0.21* | *0.25* | *1.32* | *1.81* |
| Warwickshire | Birmingham | *2.17* | *2.12* | 1.13 | 1.11 | 0.87 | 0.80 | 0.09 | 0.03 | *0.61* | *0.30* | *1.65* | *6.35* |
| SOUTH-WALES | | | | | | | | | | | | | |
| Glam & Monmouth | Cardiff | 3.44 | 2.17 | 0.32 | 0.34 | 0.82 | 0.73 | *1.00* | *0.68* | 0.10 | 0.13 | 0.48 | 0.54 |
| SOUTH | | | | | | | | | | | | | |
| London | London | 0.48 | 0.38 | *2.82* | *2.19* | *1.46* | *1.37* | 0.78 | 0.17 | 0.25 | 0.26 | 1.14 | 0.84 |

Lee (1979). LQ>1: relative high specialisation; LQ=1: no specialisation; LQ<1: relative lack of employment.

Among places in Lancashire and Yorkshire, which had been at the forefront of the industrial revolution, although both counties had a large concentration of local employment in textile manufacturing while services and new growth sectors were under-represented (Table 2.9)[13], there is a contrast between the West Riding towns of Bradford, Halifax, Leeds and Sheffield, with high *but* diminishing specialisation in textile production, which are found to be resilient, and the shock-prone Lancashire towns of Blackburn, Bolton and Wigan. It should also be said that Lancashire specialised in cotton textiles and Yorkshire in woolen textiles, and that the 1861-63 American Cotton Famine had a major adverse impact on the Lancashire cotton district. Some mills were able to diversify, for example into hat production, others could adjust to different sources of cotton, and others both spun and wove cotton, so these were able to better adapt to the shortage of raw cotton. The absence of such characteristics in some towns may have well carried over to reduce resilience, thus providing a significant explanation for their relative vulnerability. In contrast, Manchester is seen to be super-resilient, one possible reason being that its economy was more similar to the diversified towns in the Midlands than to the smaller Lancashire cotton towns, and had all the professional, administrative and commercial activities that are typically present in principal cities (Hunt, 1973).

I also find a contrast among ports towns specifically Greenock, Newcastle and Sunderland, characterised by extreme *and* increasing specialisation in shipbuilding, which are found to be shock-prone, and Cardiff and Hull. The former towns may have been subject to *lock-in effects* as suggested by evolutionary economic geography and path dependence theory. These concepts usually apply to places dominated by heavy industries, and that become locked into outmoded structures because positive externalities initially arising from clustering and specialisation then turn into growing rigidity and diminishing returns. Thus, by holding back innovation, lock-in undermines the adaptability of local economies and their ability to withstand shocks.

[13] While it is apparent that metal manufacturing was an additional leading occupation in West Riding but not so much in Lancashire, it is true that there were extensive iron works also in Blackburn, Bolton and Wigan, and both regions had railway and engineering works, paper mills, chemical plants and breweries (Great Britain Historical GIS Project, 2004).

## 2.8 Concluding remarks

In this paper I have analysed the relative resilience of British towns to recessionary shocks. The sample period covers four economic crises that occurred in 1874, 1883, 1890 and 1901, and this study has aimed to assess the actual evolution of post-1890 wages against the trajectories they are expected to have followed if they had been resilient to the 1890 and subsequent shocks. While U.K. studies of regional resilience already exist which are based on conventional econometric impulse response models, the present paper is the first to use time-series techniques in combination with state-of-the-art spatial panel econometric methods to model the impact and transmission of shocks, looking explicitly at the historical past.

To characterise the different reactions of different places to the shocks of interest, I have drawn on the resilience literature and distinguished economies on the basis of their ability to (a) resume their counterfactual wage paths by the end of the post-1890 period, a notion known as 'ecological resilience', or outperformed them thus showing positive hysteretic effects (Figure 2.7), and (b) resist to and recover from shocks leading to non-negative net differences, a notion known as 'engineering resilience' (Figure 2.8).

The main finding is that excessive and increasing specialisation in specific sectors (cotton, shipbuilding) is associated with a high level of vulnerability to recessionary shocks. This can be construed as a casual effect going from extant industrial structure to wage impact, with industrial specialisation evolving only slowly as a lagged response to the severity of economic shocks in different towns. In my view, the scope for re-orientating institutions, resources, technologies and skills in response to a shock is lower in highly specialised economies, so the ability to preserve the pre-existing growth trajectory or move to a superior one is also lower. By contrast, the flexibility stemming from a broad manufacturing and non-manufacturing base is a source of adaptability and resilience.

# Chapter 3

# Is the Wage Curve Truly an ‘Empirical Law of Economics’?
# A Spatial Panel Approach to Non-Nested Model Comparison and Selection with reference to Great Britain’s Local Areas

## 3.1 Introduction

### *3.1.1 Background*

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 estimation methods; the most prominent contributions are 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) and Australia (Kennedy and Borland, 2000). There is also a growing literature which pays attention to, and corrects for, spatial dependence arising from cross-region interactions/ spillovers, both in the German context (Buettner, 1999; Baltagi, Blien and Wolf, 2000; Longhi, Nijkamp and Poot, 2006; Elhorst, Blien and Wolf, 2007) and in the context of New Zealand (Morrison, Papps 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 pay and unemployment; this suggests imperfectly competitive labour markets where firms are not wage takers but adjust the level of pay downwards as unemployment increases. However, the approach of investigators so far has been to simply attest the empirical validity of the wage curve 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. None of these authors has taken a step further to explore whether the wage curve can also be accepted as 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 direct confrontation with alternative earnings functions.

This chapter goes beyond model fitting and slope estimation to examine the relative success of the wage curve when faced by competing hypotheses of wage determination. 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 theories have strong foundations in the urban and geographical economics literature, summarised by Huriot and Thisse (2000), Fujita and Thisse (2002), Fujita, Krugman and Venables (1999) and Brakman, Garretsen and Van Marrewijk (2009a), but they propose distinct causes of pecuniary external economies of scale and spatial economic agglomeration.

Urban Economics attributes a primary role to market linkages at *intra-regional level*; wages increase with density because of intermediate producer service linkages (assuming that final goods and services producers prefer variety, there are efficiency gains / cost advatanges from a greater variety of imperfectly-substitutable, increasing-returns intermediate market services inputs). By contrast, New Economic Geography emphasises market linkages at *inter-regional level*; wages increase with market potential because of productivity advantages from good access to large supplier and consumer markets, with higher wages associated with higher income/

demand in surrounding regions alongside lower transport costs to those regions. Hence, Urban Economics and New Economic Geography have distinct views on how economic geography can shape the spatial wages distribution, respectively looking at within- and between-region market linkages, but there is no reference to unemployment in either theory.

I use panel data for Great Britain's 408 unitary authority and local authority areas (UALADs) over the period 1999-2009, and adopt an approach to non-nested hypotheses testing which involves estimating an Artificial Nesting Model (ANM) that encompasses both rival models (Davidson and MacKinnon, 1993; Hendry, 1995). In doing so, I build 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 EU NUTS2 regions. While Fingleton (2006, 2007) assumes non-spatial errors, I allow for spatial effects working through the error term (by specifying a spatially autoregressive error components process) as well as through the spatially lagged dependent variable, following the burgeoning spatial panel econometrics literature pioneered by Kapoor, Kelejian and Prucha (2007) and Fingleton (2008a).

The chapter is organised as follows. In section 3.2 I review the relevant wage curve literature and define the model to be estimated empirically. Section 3.3 is concerned with the theoretical relationships coming from UE and NEG and the extended empirical specifications. Section 3.4 describes the variables and the data. Section 3.5 briefly sets out the methodology for estimation and instrumentation, while results are presented in section 3.6 for the wage curve, UE and NEG models in isolation. In section 3.7 the issue of comparing non-nested rival models is considered, and ANM results discussed. Section 3.8 summarises and concludes.

## 3.2 The Wage Curve model

### *3.2.1 Review of the Wage Curve literature*

The extensive literature on the wage curve which has developed in the last couple of decades postulates that, holding 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 employees 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 countries and time periods, of approximately -0.10. This empirical regularity is also documented in Blanchflower and Oswald (1995), while in Blanchflower and Oswald (2005) they confirm the existence of the US wage curve using more recent American data.

For the British case (see also Blanchflower and Oswald, 1994b), they analyse data for approximately 175,000 workers from the General Household Surveys of 1973-1990, and estimate a slope of around -0.10 after controlling for region-fixed effects and individual-worker characteristics; 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 saw some variation in wage curve elasticity across different categories. For instance, Card (1995), Baltagi and Blien (1998) and Baltagi, Blien and Wolf (2009) found that wages are more responsive to unemployment (hence the wage curve is more elastic) for men (see also Fingleton and Longhi, 2013) who tend to be employed in sectors with higher entry and exit costs as opposed to women, and also for the young, unskilled, foreigners and private sector workers all of whom tend to have weaker bargaining power. However there are some exceptions; for example, Baltagi, Blien and Wolf (2000) and Kennedy and Borland (2000) reported for Eastern Germany and Australia respectively that 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 stronger because, in depressed labour markets, they have more difficulty than others finding an alternative job when threatened by dismissal (e.g. in the event of an industrial dispute) or by lay-off (e.g. during an economic downturn), therefore employers do not need to remunerate them so well.

There are various non-competitive labour market explanations for the existence of a wage curve. The most common hypothesis is the efficiency wage or labour turnover costs theory; the argument is that, at times of higher unemployment, firms face lower costs of replacing workers while the costs to employees of job losses or voluntary quits are higher, therefore the wage that firms need to pay in excess of market-clearing to motivate workers is lower. By contrast, at times of lower unemployment, there are more jobs available to workers, as a consequence employers must offer higher wages to retain workers or to avoid shirking. Another hypothesis is 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, and this could lead to accepting a lower negotiated level of pay; nevertheless, this theory may be contradicted to the extent that unionised wages are less sensitive to the business cycle as they are set in multi-year contracts.[14]

In the GB context, one of the main contributions is that of Bell, Nickell and Quintini (2002). Their analysis is based on *British New Earnings Survey* wage data and *Labour Market Trends*/ *UK Regional Trends* unemployment data for the GB NUTS1 regions over the period 1976-1997. The wage series is compositionally corrected in order to eliminate grouped data bias (Moulton, 1990)[15]. The correction is implemented in two stages; first they regress individual-level wages on (time-varying) individual-worker characteristics, individual-fixed effects, and region-

[14] Another theory is the labour-contract model, which assumes that firms and workers maximise joint utility, meaning that higher wages may be associated with higher contractual employment and therefore with lower unemployment rates. Sato (2000) proposes a search model where equilibrium wage and unemployment levels are driven by productivity differentials across local labour markets, with higher productivity associated with higher wages and lower unemployment.

[15] The grouped nature of the sample data is a possible cause of positive cross-section dependence in the error term and downward bias in the standard error of the unemployment coefficient. 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 because individuals within the same region can be expected to share the same characteristics.

specific time dummies. The fitted values from the first-stage regression, which are treated as compositionally-corrected wages having been adjusted for individual-level variables, enter 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, with the local unemployment rate instrumented by its first-order and second-order lags. Using this approach, Bell, Nickell and Quintini (2002) obtain a (short-run) wage curve elasticity of -0.03.

Johnes (2007) accommodates grouped data effects, as well as cross-section heterogeneity, by simultaneously allowing for region- and individual-level random effects. Using *British Household Panel Survey* wage data and *Labour Force Survey* unemployment data for the period 1992-2003, and instrumenting the unemployment rate by its one-year lag, he estimates a slope of -0.05.

Hence this evidence for Britain suggests that, while a British wage curve exists, over the short term its magnitude is lower, in absolute terms, than the -0.10 'law'.

### *3.2.2 The empirical Wage Curve specification*

The wage curve literature reviewed in section 3.2.1 neglects to explicitly model cross-region interactions/ spillovers; wages in one area are described as a function of local unemployment alone, while the impact of wages and unemployment in areas nearby is ignored.

There are, however, wage curve studies which do pay attention to the implications of spatial effects for the validity and strength of Blanchflower and Oswald's 'law', and I draw from these in order to motivate the empirical specification of the wage curve model in this chapter.

One argument in favour of a spatial approach to the wage curve estimation is the presence of spatial dependence in regional unemployment; 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 various suggested reasons as to why it may occur. First, people commute between residence and workplace within functional labour market areas, however the formal/ administrative units used for data collection often do not reflect these functional areas but cut across them, so that each functional area may extend over several formal/ administrative areas. Concentrations of high (low) unemployment may also be due to the spatial patterns of population/ employment growth (labour demand), the spatial distribution of personal or job characteristics (labour supply), and the geography of house prices as disadvantaged workers tend to seek cheaper accommodation. If such effects are significant but they 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, Nijkamp and Poot (2006).

Longhi, Nijkamp and Poot (2006) estimate a wage curve for Western Germany by fixed-effects 2SLS using data for 327 regions over the period 1990-1997. Their aggregate analysis shows that the spatial lag of the unemployment rate is statistically significant but (unexpectedly) positively signed. However its inclusion enables the authors to uncover a range of spatial processes which are consistent with efficiency wage and labour turnover costs theories (monopsonistic competition in local labour markets). The coefficient on the interaction between local unemployment and its spatial lag is significant and negative, meaning that the wage curve is more elastic (i.e. the negative response of wages to local unemployment is stronger, the unemployment elasticity of pay is more negative) if employment opportunities in surrounding regions are tighter; this indicates that employers are more likely to reduce wages in response to an increase in local unemployment without fearing that workers will quit their current jobs for similar ones in adjacent areas.

They also show a positive and significant coefficient on the interaction between local unemployment and local agglomeration/ accessibility, suggesting that relatively remote, isolated areas have a more elastic wage curve since mobility costs associated with a job change (commuting, migration, job search) are higher. These effects are 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.

In Buettner (1999), using data for 327 regions of Western Germany over the period 1987-1994, the unemployment rate in neighbouring regions is shown to have a negative as well as significant effect on local pay. There is also a noticeable reduction in the estimated coefficient on local unemployment once its spatially lag is included, possibly because the presence of spatial dependence makes local unemployment a proxy for regional as well as local wage curve effects.

Another source of model misspecification and incorrect (here inconsistent as well as biased) estimates is spatial correlation in regional wages. Buettner (1999) tests for this type of spatial effects and finds strong support for the hypothesis that wage rates in neighbouring districts exert an autonomous influence on local pay. Despite the significance of the spatially lagged dependent variable, the change in the estimated coefficients on the explanatory variables is negligible, possibly because the endogenous spatial lag picks up spatial effects which are only weakly correlated with the model's regressors.

A further type of cross-section dependence involves the error term, and typically reflects common shocks or 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 residuals correlation leads to reduced standard errors, inflated *t*-ratios and unreliable inference. Existing wage curve studies correct for spatial residuals dependence by means of a common factor approach. For example, in the context of Eastern Germany, Baltagi, Blien and Wolf (2000) use 2SLS estimation on variables in first differences as a way to wipe out region-fixed effects – e.g. relating to a common history or locally available natural resources - which may cause dependency among closely located areas. As an alternative strategy, which is equivalent to the inclusion of time-fixed effects, Pesaran (2006) suggests estimating an augmented model with cross-section averages of the regressand and regressors acting as placeholders for unobserved common factors. Elhorst, Blien and Wolf (2007) introduce spatial first-difference 2SLS and eliminate both region-fixed and time-fixed effects both by differencing in time and by taking the value of the dependent and independent variables for each region in deviation from that in one reference region. Therefore, these authors do not explicitly define spatial effects via weights matrices as typically invoked in spatial econometrics. Likewise, in Buettner (1999) and Longhi et al. (2006), empirical specifications with a spatial lag of the error term are not considered, which they justify by a non-significant Lagrange Multiplier test of spatial residuals autocorrelation.

This chapter follows the strand of the wage curve literature which deals with spatial effects but also rethinks the error term structure. Blanchflower and Oswald's relationship is thus extended by adding a spatially autoregressive error process to a random-effects specification, assuming autocorrelated errors in space and time, together with spatial lags of the dependent and independent variables. The different spatial mechanisms can be summarised as follows

$$\begin{aligned}\mathbf{Log}\boldsymbol{Wage} &= b_0 + \rho(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{Wage} + b_1\mathbf{Log}\boldsymbol{U} + \theta(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{U} + b_2\mathbf{Log}\boldsymbol{T} + b_3\mathbf{Log}\boldsymbol{S} + \\ &\quad + b_4\mathbf{Log}\boldsymbol{A} + \omega_1(\mathbf{Log}\boldsymbol{U})((\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{Wage}) + \omega_2(\mathbf{Log}\boldsymbol{U})(\mathbf{Log}\boldsymbol{A}) + \mathbf{e} \\ \mathbf{e} &= \lambda(\mathbf{I}_T \otimes \mathbf{M})\mathbf{e} + \boldsymbol{\xi} \\ \boldsymbol{\xi} &= (\boldsymbol{\iota}_T \otimes \mathbf{u}) + \mathbf{v}\end{aligned} \tag{3.1}$$

where the $N$ x 1 vector $\mathbf{e}$ is the spatially dependent error term, and this is a function of innovations $\boldsymbol{\xi}$ which combine a permanent random-effects component $\mathbf{u} \sim iid(0, \sigma_u^2)$, and a transient remainder or disturbance component $\mathbf{v} \sim iid(0, \sigma_v^2)$. $\mathbf{W}$ and $\mathbf{M}$ are $N$ x $N$ spatial weights matrices, defining the structure and intensity of interactions and spillovers among areas; these are specified in section 3.4.

The variable $\mathbf{Log}\boldsymbol{A}$ refers to an index of accessibility/ agglomeration, and is computed following Longhi et al. (2006) as the sum of total employment $Emp$ in neighbouring districts weighted by distance $d$, i.e. $A_{i,t} = \sum_j kEmp_{j,t} d_{ij}^{-1}$ [16] [17].

A set of additional explanatory variables measuring local labour efficiency ( $\boldsymbol{T}$ and $\boldsymbol{S}$ ), as described in section 3.4, is also added on the right-hand side of the wage curve to control for composition effects. This is in line with applied wage curve studies using micro data which often include regional averages of individual-level

[16] $k$ is a scaling parameter, here equal to $10^{-6}$.

[17] Following Longhi et al. (2006) I 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 other regions with which region $i$ could potentially interact, as in $A_{i,t} = \sum_j k(E_{j,t} E_{i,t})^{0.5} d_{ij}^{-1}$ . This spatial interaction measure gives an identical set of estimation results so I do not report these.

attributes (e.g. skills level, social deprivation). If the monopsonistic power of employers in areas with a skilled and qualified workforce is lower as one would anticipate, then I should expect wages to increase with $\boldsymbol{T}$ ($b_2$>0) and to decrease with $\boldsymbol{S}$ ($b_3$<0).

## 3.3 The rival Urban Economics (UE) and New Economic Geography (NEG) models

### *3.3.1 The model motivated by 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)[18]. I estimate a relationship linking wage/productivity levels and density of economic activity, in the form of employment density, and allowing for a direct test of the existence of increasing returns to economic agglomeration. Local production conditions, particularly the varying regional strength of intermediate producer services lunkages (which increases with economic density), 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 constant returns, and a non-traded intermediate market services sector (*I*), characterised by monopolistic competition and (internal) increasing returns, which provides inputs to the output (*Q*) of competitive industry *M*. So, assuming a Cobb-Douglas production function for *M*, I have

$$Q = (M^{\beta} I^{1-\beta})^{\alpha} Land^{1-\alpha} = (M^{\beta} I^{1-\beta})^{\alpha} \tag{3.2}$$

where land (*Land*) is equal to one because production is per unit area or squared km, *M* is the level of labour efficiency units employed in making *M*'s goods and services directly, and *I* is the level of intermediate market services inputs based on a CES (Constant Elasticity of Substitution) sub-production function under product variety,

[18] There are alternative UE models, e.g. Combes, Mayer and Thisse (2008) and Brakman, Garretsen and Van Marrewijk (2009b), which lead to the same reduced form.

monopolistic competition and internal scale economies *à la* Dixit and Stiglitz (1977) that is [19]

$$I = \left[ \sum_{z=0}^{x} i_z^{1/\mu} \right]^{\mu} \tag{3.3}$$

As shown in Appendix 3, this leads to a relationship between the level of final goods and services output per unit area (*Q*) and the level of total effective labour (in both the final goods and services sector and the intermediate market services sector) per unit area (*L*), as in

$$Q = (M^{\beta} I^{1-\beta})^{\alpha} = \phi L^{\gamma} \tag{3.4}$$

with constants $\phi$ and elasticity $\gamma$ where

$$\gamma = \alpha[1 + (1-\beta)(\mu - 1)] \tag{3.5}$$

*L* is given by the product between total employment level per unit area (*E*) and each area's labour efficiency level (*H*), i.e. $L=E \cdot H$. Parameter $\mu$ ($\mu>1$) 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. Whether or not there are external increasing returns ($\gamma>1$) depends on the amount of internal scale economies being sufficiently large ($\mu>1$); on the non-traded, increasing-returns sector being sufficiently important to final production (which is indexed by the magnitude of $\beta<1$); and on diminishing returns due to congestions costs[20] ($1-\alpha<1$) being small enough so as not to outweigh the other two factors.

Thus, denser districts tend to have (more than proportionally) higher productivity and wage levels, indicating external increasing returns to scale (external scale economies), because final goods and services producers have a preference for a

---

[19] The *I* sector comprises **producer services**, or support activities with high-information content, e.g. banking, finance, insurance, real estate, business and other professional services; they sometimes reflect a "contracting out" of services that could be provided in-house (OECD). These can be considered as being characterised by low entry and exit costs and by many small firms producing highly-differentiated services with appreciable internal scale economies e.g. due to fixed costs associated with the business start-up, which is close to what is implied by monopolistic competition.

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

greater variety, common in larger towns and cities, of imperfectly-substitutable, increasing-returns intermediate market services inputs. Hence, efficiency gains / cost advantages from internal increasing returns to scale in the *I* sector translate into external scale economies (productivity gains) in the competitive, constant-returns final goods and services sector *M*.

To determine the wage rate per labour efficiency unit ($Wage$), I use the equilibrium allocation of labour inputs to final production $Q$. This entails calculating the derivative of $Q = [f(L)]^{\alpha} Land^{1-\alpha}$ with respect to $L$, or marginal product of effective labour

$$\begin{aligned} \mathrm{d}Q/\mathrm{d}L &= Land^{1-\alpha} \cdot \alpha \cdot f(L)^{\alpha} \cdot f(L)^{-1} \\ &= f(L)^{\alpha} Land^{1-\alpha} \cdot \alpha / f(L) \\ &= \alpha Q / f(L) \end{aligned} \tag{3.6}$$

then replacing $f(L)$ with $L$ without losing meaning and finally, following standard competitive equilibrium theory, setting $Wage$ equal to the marginal product of $L$

$$Wage = \alpha Q / L \tag{3.7}$$

where coefficient $\alpha$ (eq. 3.2), is the share of final output going to effective labour. Taking the natural logarithm of both sides of eq. (3.7) gives

$$\mathbf{Log}\boldsymbol{Wage} = \mathbf{Log}\boldsymbol{Q} + Log(\alpha) - \mathbf{Log}\boldsymbol{L} \tag{3.8}$$

and substituting for $Q = \phi L^{\gamma}$ and for $L = E \cdot H$ gives the short-run wage equation

$$\begin{aligned} \mathbf{Log}\boldsymbol{Wage} &= Log(\phi) + \gamma \mathbf{Log}(\boldsymbol{H} \cdot \boldsymbol{E}) + Log(\alpha) - \mathbf{Log}(\boldsymbol{H} \cdot \boldsymbol{E}) \\ &= k_1 + (\gamma - 1)\mathbf{Log}\boldsymbol{E} + (\gamma - 1)\mathbf{Log}\boldsymbol{H} \end{aligned} \tag{3.9}$$

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

### *3.3.2 The model motivated by NEG theory*

This section sets out the NEG model, following Harris (1954) and Krugman (1991). The concept of market potential dates back to Harris (1954), but Krugman (1991) was the first to develop a structural model around Harris's (1954) initial formulation, using the theory of product variety, monopolistic competition and internal scale economies introduced by Dixit and Stiglitz (1977), based on a CES (Constant Elasticity of Substitution) sub-production function; this theory, with its utility / profit maximising microfoundations and market structure assumptions which give rise to pecuniary externalities, provided the theoretical justification for the observed economic agglomeration phenomenon.

Under Krugman's (1991) general equilibrium model (i.e. the basic NEG specification), 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 mean 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 neighbouring markets. Thus the <u>wage rate per worker</u> is

$$WageL_{i,t} = \frac{\overline{Wage}_{i,t}}{Labour_{i,t}} = \left[\sum_j Y_{j,t}(G_{j,t})^{\varepsilon-1}(Tr_{ij})^{1-\varepsilon}\right]^{\frac{1}{\varepsilon}} = \left[\sum_j Y_{j,t}(G_{j,t})^{\varepsilon-1}\frac{1}{(Tr_{ij})^{\varepsilon-1}}\right]^{\frac{1}{\varepsilon}} = MP_{i,t}^{\frac{1}{\varepsilon}}$$

where $\overline{Wage}_{i,t}$ denotes area *i*'s total wage bill and $Labour_{i,t}$ denotes area *i*'s total effective labour. The <u>wage rate per labour efficiency unit</u> is

$$\begin{aligned} Wage_{i,t} &= \frac{\overline{Wage}_{i,t}}{Emp_{i,t} \cdot H_{i,t}} = MP_{i,t}^{\frac{1}{\varepsilon}} \\ Wage_{i,t} &= \frac{\overline{Wage}_{i,t}}{Emp_{i,t}} = MP_{i,t}^{\frac{1}{\varepsilon}} H_{i,t} \end{aligned} \qquad (3.10a)$$

where $Emp_{i,t}$ denotes area *i*'s total employment level and $H_{i,t}$ is its labour efficiency level. Taking logs and using matrix notation, I have

$$\textbf{Log}\boldsymbol{Wage} = \frac{1}{\varepsilon}\textbf{Log}\boldsymbol{MP} + \textbf{Log}\boldsymbol{H} \tag{3.10b}$$

Eq. (3.10b) states that wages in region $i$ at time $t$ have a propensity to be higher if income ($Y$) in neighbouring markets is higher, as this would raise demand for region $i$'s goods. Since the elasticity of substitution among varieties is greater than one (ε>1), wages tend to be higher if distance of $i$ from those markets and thus transport costs ($Tr$) are lower. Also, since ε>1, wages are boosted by a higher price index ($G$) in neighbouring markets, as this would indicate that there are less varieties sold in these markets and so less competition facing region $i$'s firms.

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

$$MP_{i,t} = \sum_j Y_{j,t} d_{ij}^{-1} \tag{3.11}$$

which is actually the same as Krugman's (1991) structural model with the assumptions that $G$ =1 (constant regional price indices, i.e. nominal market potential measure) and $\varepsilon$ =2 (it must be $\varepsilon$ >1 at the least).

Therefore, following Head and Mayer (2004), Combes, Duranton and Gobillon (2008), Mion and Naticchioni (2009) and Brakman, Garretsen and Van Marrewijk (2009b), I 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. In addition, Brakman, Garretsen and Van Marrewijk (2009b) show that market potential as measured by distance-weighted population gives identical outcomes as market potential in terms of real income/purchasing power. The illustration in Appendix 5 confirms that this is a good approximation.

Below I discuss the validity of assumptions behind the definition of market potential suggested by Harris (1954).

- $G$ =1 (prices are constant across regions, or nominal market potential measure) – this assumption is driven by data availability, in that price indices are available for the UK as a whole but not for its constituent regions, however any measurement bias is corrected for by the use of instrumental variables.
- $\varepsilon$ =2 – this parameter value satisfies the love for variety and imperfect substitution requirement in the theory of Dixit and Stiglitz (1977) that the constant elasticity of substitution among varieties must be at least greater than one.
- Internal transport costs are equal to unity – Krugman's (1991) structural model makes the same assumption, moreover there will be no serious consequences in terms of measurement bias in market potential as long as within-region transport costs are smaller than between-region transport costs, which is a reasonable assumption.
- Transport costs as function of straight-line distance – this assumption would be inappropriate for countries such as Japan with a mountainous landscape, potentially affecting the transportation route network, and where the regional hubs are mostly located along the coastline. Given the topography of the UK, it is more realistic to assume that the majority of economic activity occurs in the geometric centre of each region, and so direct distance between centres seems to be a good measure of transport costs.

### *3.3.3 The empirical UE and NEG model*

*3.3.3.1 Assumptions about labour efficiency*

This section is concerned with the measurement of effective labour ( $\boldsymbol{H}$ ) and the modelling of spatial effects (commuting effects). First I assume that $\boldsymbol{H}$ is affected by differences between workers in their ability to make productive use of the available technology; technology itself is assumed to be homogenous across areas. I thus express each area's level of labour 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, I focus on the percentage of working-age resident population with no qualifications. I denote this variable by $\boldsymbol{S}$. Another determinant of local labour efficiency is the size of the local knowledge base. This variable is denoted by $\boldsymbol{T}$, and is approximated by the relative concentration of local employment in technology-intensive computing and R&D sectors. A more precise description of $\boldsymbol{S}$ and $\boldsymbol{T}$ is given in section 3.4.

I also recognise that workers are mobile, and wages paid at the workplace depend on the level of labour efficiency 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 $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{H}$, which represents the matrix product of $(\mathbf{I}_T \otimes \mathbf{W})$ (where $\mathbf{W}$ is specified in eq. 3.19 in section 3.4) and $\boldsymbol{LogH}$; more precisely, the contribution to region *i*'s labour efficiency from in-commuting is given by element *i* of vector $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{H}$, which contains the weighted sum of labour efficiency in all other designated areas.

Combining the (exogenous) variables assumed to determine the level of local effective labour, I have

$$\begin{aligned} &\mathbf{Log}\boldsymbol{H} = c_0 + c_2\mathbf{Log}\boldsymbol{T} + c_3\mathbf{Log}\boldsymbol{S} + \rho(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{H} + \boldsymbol{\vartheta} \\ &\boldsymbol{\vartheta} \sim iid(0, \sigma_\vartheta^2) \end{aligned} \tag{3.12a}$$

which, taking just one cross-sectional regression at time *t* and assuming $|\rho| < 1$, can be rewritten as

$$\begin{aligned} \mathbf{Log}\boldsymbol{H}_t &= c_0 + c_2\mathbf{Log}\boldsymbol{T}_t + c_3\mathbf{Log}\boldsymbol{S}_t + \rho\mathbf{WLog}\boldsymbol{H}_t + \boldsymbol{\vartheta}_t = \\ &= (\mathbf{I}_N - \rho\mathbf{W})^{-1}(c_0 + c_2\mathbf{Log}\boldsymbol{T}_t + c_3\mathbf{Log}\boldsymbol{S}_t + \boldsymbol{\vartheta}_t) \end{aligned} \tag{3.12b}$$

The presence of Leontief Inverse matrix $(\mathbf{I}_N - \rho\mathbf{W})^{-1}$ means the model captures the totality of the effects influencing local labour efficiency, i.e. not only $\boldsymbol{S}$ and $\boldsymbol{T}$ locally but also the effects of $\boldsymbol{S}$ and $\boldsymbol{T}$ in other areas together with the local and remote effects of unmodelled factors. In other words, the level of labour efficiency in region *i* depends on schooling, technology and the shocks within the area ($\mathbf{Log}\boldsymbol{T}$, $\mathbf{Log}\boldsymbol{S}$ and $\boldsymbol{\vartheta}$), in neighbouring areas ($(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{T}$, $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{S}$ and $(\mathbf{I}_T \otimes \mathbf{W})\boldsymbol{\vartheta}$), in areas which are neighbours of the neighbours ($(\mathbf{I}_T \otimes \mathbf{W})^2\mathbf{Log}\boldsymbol{T}$, $(\mathbf{I}_T \otimes \mathbf{W})^2\mathbf{Log}\boldsymbol{S}$ and $(\mathbf{I}_T \otimes \mathbf{W})^2\boldsymbol{\vartheta}$), and so on with effects eventually feeding back to *i*.

Given the definition of the elements of $\mathbf{W}$ in eq. 3.19 as $W_{ij}^* = \exp(-\hat{\tau}_i d_{ij})$ (i.e. exponential inverse decay function of distance) and provided that $\rho > 0$, region *i*'s labour efficiency level will be mainly influenced by the levels of $\boldsymbol{S}$, $\boldsymbol{T}$ and $\boldsymbol{\vartheta}$ in nearby areas (distance $d_{ij}$ is small); also, this indirect effect will be higher if the weight carried by surrounding areas, which by construction falls ever more steeply as distance from *i* increases, does so less rapidly (distance decay rate $\hat{\tau}_i$ is smaller).

It should be noted that my indicators of local labour efficiency, namely $\mathbf{Log}\boldsymbol{T}$ and $\mathbf{Log}\boldsymbol{S}$, may pick up non-pecuniary externalities operating through non-market interactions and depending on the technological and skill content of local employment. The appropriate sign and statistical significance of the coefficients on $\mathbf{Log}\boldsymbol{T}$ and $\mathbf{Log}\boldsymbol{S}$ might thus indicate the presence of 'technological externalities' involving localised knowledge spillovers[21].

[21] 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, 1920) or of innovative activity (Arrow, 1962).

*3.3.3.2. Derivation of the extended specifications*

- For UE, to find $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{H}$ in terms of known variables, I then rearrange eq. (3.9) into $\mathbf{Log}\boldsymbol{H} = -k_1/(\gamma-1)+1/(\gamma-1)\,\mathbf{Log}\boldsymbol{Wage}-\mathbf{Log}\boldsymbol{E}$ and multiply both sides by $(\mathbf{I}_T \otimes \mathbf{W})$, giving

$$(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{H} = (\mathbf{I}_T \otimes \mathbf{W})\frac{-k_1}{\gamma-1}+\frac{1}{\gamma-1}(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{Wage}-(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{E} \quad (3.13)$$

Substituting the expressions for $\mathbf{Log}\boldsymbol{H}$ (eq. 3.12a) and $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{H}$ (eq. 3.13) into $\mathbf{Log}\boldsymbol{Wage} = k_1+(\gamma-1)\mathbf{Log}\boldsymbol{E}+(\gamma-1)\mathbf{Log}\boldsymbol{H}$, and adding error term $\boldsymbol{\varsigma}$, I obtain

$$\begin{aligned}\mathbf{Log}\boldsymbol{Wage} &= k_1+(\gamma-1)\mathbf{Log}\boldsymbol{E}+(\gamma-1)\Big[c_0+c_2\mathbf{Log}\boldsymbol{T}+c_3\mathbf{Log}\boldsymbol{S}+\\ &+\rho\left((\mathbf{I}_T \otimes \mathbf{W})\frac{-k_1}{\gamma-1}+\frac{1}{\gamma-1}(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{Wage}-(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{E}\right)+\boldsymbol{\vartheta}\Big]+\boldsymbol{\varsigma}\end{aligned} \quad (3.14)$$

Hence, simplifying I have

$$\begin{aligned}\mathbf{Log}\boldsymbol{Wage} &= a_0+\rho(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{Wage}+(\gamma-1)(\mathbf{Log}\boldsymbol{E}\text{-}\rho(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{E})+\\ &\quad +a_2\mathbf{Log}\boldsymbol{T}+a_3\mathbf{Log}\boldsymbol{S}+\mathbf{e}\\ \mathbf{e} &= \lambda(\mathbf{I}_T \otimes \mathbf{M})\mathbf{e}+\boldsymbol{\xi}\\ \boldsymbol{\xi} &= (\boldsymbol{\iota}_T \otimes \mathbf{u})+\mathbf{v}\end{aligned} \quad (3.15)$$

where $a_0$ is a constant, $(\mathbf{I}_T \otimes \mathbf{W})k_1$ can be ignored without effect, and the error term has the same spatial structure as that of the empirical wage curve model in eq. (3.1).

- Likewise for NEG, to find $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{H}$ in terms of known variables, I rearrange eq. (3.10b) into $\mathbf{Log}\boldsymbol{H} = \mathbf{Log}\boldsymbol{Wage} - \frac{1}{\varepsilon}\mathbf{Log}\boldsymbol{MP}$ and multiply both sides by $(\mathbf{I}_T \otimes \mathbf{W})$, giving

$$(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{H} = (\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{Wage}-\frac{1}{\varepsilon}(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{MP} \quad (3.16)$$

Substituting the expressions for $\mathbf{Log}\boldsymbol{H}$ (eq. 3.12a) and $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{H}$ (eq. 3.16) into $\mathbf{Log}\boldsymbol{Wage} = \frac{1}{\varepsilon}\mathbf{Log}\boldsymbol{MP}+\mathbf{Log}\boldsymbol{H}$, and adding error term $\boldsymbol{\varsigma}$, I obtain

$$\begin{aligned}\mathbf{Log}\boldsymbol{Wage} = \frac{1}{\varepsilon}\mathbf{Log}\boldsymbol{MP} + [c_0 + c_2\mathbf{Log}\boldsymbol{T} + c_3\mathbf{Log}\boldsymbol{S} + \\ + \rho\left( (\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{Wage} - \frac{1}{\varepsilon}(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{MP} \right)\Bigg] + \boldsymbol{\varsigma}\end{aligned} \tag{3.17}$$

Hence, simplifying, the extended empirical specification of the NEG model can be summarised as follows

$$\begin{aligned}\mathbf{Log}\boldsymbol{Wage} = a_0 + \rho(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{Wage} + \frac{1}{\varepsilon}(\mathbf{Log}\boldsymbol{MP}\text{-}\rho(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{MP}) + \\ + a_2\mathbf{Log}\boldsymbol{T} + a_3\mathbf{Log}\boldsymbol{S} + \mathbf{e} \\ \mathbf{e} = \lambda(\mathbf{I}_T \otimes \mathbf{M})\mathbf{e} + \boldsymbol{\xi} \\ \boldsymbol{\xi} = (\boldsymbol{\iota}_T \otimes \mathbf{u}) + \mathbf{v}\end{aligned} \tag{3.18}$$

where $a_0$ is a constant, and the error term has the usual spatial structure.

Other researchers have recognised the need to consider skills variation, for example Combes, Duranton and Gobillon (2008). Specifically for NEG, Head and Mayer (2006) augmented the simple wage equation of Krugman (1991) by adding labour efficiency variables to the  original NEG model.

## 3.4 Variables and data

| | Source[22] | Description | Mean | Min | Max |
|---|---|---|---|---|---|
| *Wage* | Annual Survey of Hours and Earnings (ASHE) | **Mean Gross Weekly Wage Rate (workplace-based)** (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 (residence-base)** (proportion of working-age population claiming unemployment-related benefits[23]) | 2.4% | 0.2% | 10.5% |
| *E* | Annual Business Survey (ABS) | **Employment Density (workplace-based / survey of employers)** (total employment per square kilometre) | 867 | 3 | 96,125[24] |
| *MP* | ONS mid-year population estimates | **Market Potential** (distance-weighted sum of population in adjacent areas, as in eq. 11 of Ch. 3.3.2) | 2,261 | 358 | 5,468 |

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 table above contains details on data sources, variable definitions and summary statistics for the variables in question, excluding transformations of these

[22] All data is available from NOMIS, the Office for National Statistics' on-line labour market statistics database.

[23] Since 1996 only people claiming Jobseeker's Allowance have been counted.

[24] City of London.

such as interaction terms and spatial lags. (Log) wage spans the period 2000-2010 while all predictors, except (log) $\boldsymbol{S}$ which refers to year 2001, denote $N$ x 1 vectors (for $N$=408 regions) at time $t$ ($t$=1999…2009).

The $N$ x $N$ standardised spatial weights matrix $\mathbf{W}$ for the endogenous spatial lag is a commuting-based matrix taking the form of an exponential inverse decay function of distance, as follows

$$
\begin{aligned}
W_{ij}^{*} &= \exp(-\hat{\tau}_i d_{ij}) \text{ for } i \neq j \\
W_{ij}^{*} &= 0 \qquad\qquad\quad \text{for } i = j \\
W_{ij}^{*} &= 0 \qquad\qquad\quad \text{for } \mathrm{d}_{ij} > 100\text{km} \\
W_{ij} &= \frac{W_{ij}^{*}}{\sum_{j=1}^{N} W_{ij}^{*}}
\end{aligned}
\tag{3.19}
$$

where $\hat{\tau}_i$ is calibrated on commuting flows (Fingleton, 2003) (see Appendix 4).

It should be noted that the commuting data used to obtain $\mathbf{W}$ are 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.

Moreover, weights are allotted to distances up to 100km to accommodate long-distance commuting, for example workers travelling further than 40km (which amounted to 4.3% in 1991), with areas beyond 100km given zero weight.

This $\mathbf{W}$ matrix is used to construct the following regressors:

- $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{Wage}$, the spatial lag of local wages, given by the matrix product between $\mathbf{W}$ and the (log) wage vector at each time period $t$ (the associated coefficient indicates the responsiveness of a region's wage rate to that of neighbouring regions located within commuting distance);
- $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{U}$, the spatial lag of local unemployment, given by the matrix product between $\mathbf{W}$ and the (log) unemployment vector at each time period $t$ (its

inclusion allows testing whether local wage rate variance can be explained by the unemployment rate in neighbouring regions located within commuting distance);

○ **Log*E*-**$\rho$**(**$\mathbf{I}_T \otimes \mathbf{W}$**)Log*E*** and **Log*MP*-**$\rho$**(**$\mathbf{I}_T \otimes \mathbf{W}$**)Log*MP*** , the composite employment density and market potential variables derived previously.

The spatial weights matrix **M** for the error process is a binary, 'natural' neighbours, or contiguity-based matrix, given by

$$
\begin{aligned}
&M_{ij}^* = 1 \text{ if } i \text{ and } j \text{ are contiguous i.e. share a border} \\
&M_{ij}^* = 0 \text{ otherwise} \\
&M_{ij} = \frac{M_{ij}^*}{\sum_{j=1}^{N} M_{ij}^*}
\end{aligned}
\tag{3.20}
$$

## 3.5 Methodology

The Spatial Two-Stage Least Squares / Generalised Method of Moments (S2SLS / GMM) estimation procedure is described in section 2.4 of the previous chapter and so I refer to this for details.

The instrumentation strategy is also outlined in section 2.4. More specifically, following the widely accepted and well established approach of Drukker, Egger and Prucha (2013), I specify an instruments set which includes the linearly independent columns of

$$
\begin{gathered}
\left[\mathbf{X}_f ; (\mathbf{I}_T \otimes \mathbf{W})\mathbf{X}_f ; (\mathbf{I}_T \otimes \mathbf{M})\mathbf{X}_f ; (\mathbf{I}_T \otimes \mathbf{W})(\mathbf{I}_T \otimes \mathbf{M})\mathbf{X}_f\right] \\
\mathbf{X}_f = \left[\mathbf{Log}\boldsymbol{S}_i ; \mathbf{Log}\boldsymbol{T}_{i,t-1}\right]
\end{gathered}
\tag{3.21}
$$

with spatial lags of first order as is common choice in applied spatial econometrics and in order to avoid over-parameterisation. This formulation relies on the variables in $\mathbf{X}_f$ being exogenous. Local technical knowledge (***T***) is temporally lagged by one year so that it predates the period of analysis and is predetermined with respect to **Log*Wage***, thus can be treated as exogenous. The educational attainment variable (***S***) postdates **Log*Wage*** by one year, but can be treated as exogenous for two reasons;

because it is unlikely to see feedback from **Log*Wage*** to schooling on this time scale, and because schooling is affected by factors other than wage differentials - such as administrative institutions, policy initiatives and socio-cultural differences - so that it is in effect predetermined[25].

The key regressors (unemployment rate, employment density and market potential) are temporally lagged by one year so that they predate the period of analysis and are predetermined with respect to **Log*Wage***, thus can be treated as exogenous. The spatial lag of **Log*Wage*** is clearly endogenous because of multilateral spatial dependence between wage observations and is thus identified using appropriate instruments as in eq. (3.21).

[25] Comparing the percentage of residents with no qualifications in 1991 and 2001, one would find that while this has fallen dramatically there exists a strong linear correlation (0.872) between the 1991 and 2001 datasets, so using the 1991 Census would give similar results.

## 3.6 Estimation results

### *3.6.1 The spatial Wage Curve model*

**Table 3.1 Results from the spatial Wage Curve model estimated in isolation**

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| *Treatment of* Log$U$ and its Spatial Lag | **Exogenous** | **Exogenous** | **Exogenous** | **Exogenous** | **Endogenous** |
| Spatial Spillovers/ Monopsony Effects | | | | | |
| Endog. Spatial Lag ($\rho$) | **0.0859** | **0.1068** | **0.0891** | **0.1066** | **0.1336** |
| (t-stat) | (5.11)*** | (8.07)*** | (6.64)*** | (8.17)*** | (6.52)*** |
| Local Unemployment Rate | | | | | |
| Log$U$ ($b_1$) | **-0.1122** | **0.0131** | **-0.0236** | | |
| (t-stat) | (-0.85) | (1.06) | (-2.01)*** | | |
| Spatial Lag of Log$U$ | | | | | |
| $(\mathbf{I}_T \otimes \mathbf{W})$Log$U$ ($\theta$) | **-0.1239** | **-0.1404** | | **-0.1289** | **-0.2684** |
| (t-stat) | (-4.45)*** | (-6.47)*** | | (-5.99)*** | (-3.35)*** |
| Local Tech. Knowledge Base | | | | | |
| Log$T$ ($b_2$) | **0.0344** | **0.0365** | **0.0459** | **0.0378** | **0.0320** |
| (t-stat) | (8.27)*** | (8.42)*** | (10.62)*** | (9.40)*** | (6.77)*** |
| Local Unskilled Workforce | | | | | |
| Log$S$ ($b_3$) | **-0.1160** | **-0.1375** | **-0.0857** | **-0.1146** | **-0.1168** |
| (t-stat) | (-3.92)*** | (-4.51)*** | (-2.82)*** | (-4.98)*** | (-4.72)*** |
| Local Accessibility/Agglom. | | | | | |
| Log$A$ ($b_4$) | **0.1197** | | | | |
| (t-stat) | (4.34)*** | | | | |
| Interaction of Log$U$ with its | | | | | |
| Spatial Lag ($\omega_1$) | **-0.0113** | | | | |
| (t-stat) | (-0.52) | | | | |
| Interaction of Log$U$ with | | | | | |
| Log$A$ ($\omega_2$) | **-0.0184** | | | | |
| (t-stat) | (-0.97) | | | | |
| Constant ($b_0$) | 7.1321 | 6.2648 | 6.1948 | 6.1992 | 6.0937 |
| (t-stat) | (28.47)*** | (51.39)*** | (50.43)*** | (58.30)*** | (50.96)*** |
| *Error process* | | | | | |
| λ [a] | **0.6627*** ** | **0.6681*** ** | **0.6309*** ** | **0.6919*** ** | **0.7725*** ** |
| $\sigma_v^2$ | 0.0035 | 0.0036 | 0.0043 | 0.0035 | 0.0031 |
| $\sigma_1^2 = \sigma_v^2 + T\sigma_u^2$ | 0.0561 | 0.0614 | 0.0643 | 0.0594 | 0.0646 |
| CORR [b] | **0.7419** | **0.7251** | **0.6903** | **0.7233** | **0.6845** |
| No. areas | 408 | 408 | 408 | 408 | 408 |
| No. in-sample years (1999-2009) | 11 | 11 | 11 | 11 | 11 |

[a] Significance about the spatial autoregressive coefficient is based on bootstrap inference (see Ch. 2).
[b] Correlation between observed and fitted values of **Log*Wage***.
* is significance at 10% level, ** is significance at 5% level, *** is significance at 1% level.

Table 3.1 shows results of fitting wage equations motivated by the wage curve literature. The starting point is the specification in eq. (3.1) of section 3.2.2. This model (column 1) gives an unemployment elasticity of pay which is in line with the empirical regularity of -0.10, but has a *P*-value of 0.28 (*t*-ratio -0.85) which largely exceeds conventional Type I error rates for a two-tailed significance test. Meanwhile, its spatial lag is both close to -0.10 and highly relevant. The other estimated coefficients are all appropriately signed and statistically significant except for the interaction terms. In particular, the coefficient on the interaction between $\mathbf{Log}\boldsymbol{U}$ and $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{U}$ is correctly (negatively) signed, suggesting that the wage curve is more elastic in areas surrounded by high unemployment (Longhi et al., 2006), but has no additional explanatory power; as for the interaction between $\mathbf{Log}\boldsymbol{U}$ and $\mathbf{Log}\boldsymbol{A}$, not only is this statistically insignificant but it is also wrongly (negatively) signed in the light of the expectation (Longhi et al., 2006) that more accessible/agglomerated areas should have a less elastic wage curve.

There is strong evidence of a significant spatially autoregressive process involving $\mathbf{Log}\boldsymbol{Wage}$ and the error term. As suggested by $\hat{\rho}$ = 0.0859 with a *t*-ratio of 5.11, wages within commuting distance have a positive and significant effect on local pay; in addition to spatial spillovers from commuting links, what may be happening here is that higher wages in nearby areas raise the opportunity wage of local workers, thereby increasing the wage that employers must pay to attract or retain workers. Hence, while $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{Wage}$ in the urban economics and economic geography models comes from an auxiliary SAR process involving labour efficiency (section 3.3.3), here it represents something different i.e. spatial spillovers and/ or monopsonistic competition in local labour markets. The positive and significant spatial autoregressive error term parameter ( $\hat{\lambda}$ =0.6627) points to the effect of global shocks transmitted across the urban hierarchy or to the presence of positive spatial autocorrelation in unobserved/ unmeasured causes of interregional heterogeneity.

Removing the endogenous spatial lag (unreported results) would make the coefficient on $\mathbf{Log}\boldsymbol{U}$ significant but its size would become unreasonably large, in absolute terms, compared to the -0.10 'law'; therefore the spatially lagged dependent variable should be in the model. By contrast, one could argue that $\mathbf{Log}\boldsymbol{A}$ should be removed because it might be capturing the same proximity effects as measured by market potential, and I aim to maintain a clear distinction between the wage curve and its rival (non-nested) NEG model. Moreover, as seen above, there is no evidence of significant interaction variables.

This leads to the second (and subsequent) models without $\mathbf{Log}\boldsymbol{A}$ and the interaction terms, which reaffirm the finding that local wages mainly respond to unemployment within commuting distance. In Model 2, the coefficient on $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{U}$ remains close to -0.10 and strongly significant, whereas the coefficient on $\mathbf{Log}\boldsymbol{U}$ is now positive as well as insignificant. Further evidence is provided by Models 3 and 4; by removing $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{U}$ (Model 3), the coefficient on $\mathbf{Log}\boldsymbol{U}$ becomes appropriately signed and statistically significant but its absolute value is well below the expected size (although close to UK estimates in Bell, Nickell and Quintini, 2002 and Johnes, 2007), whereas an empirical specification with $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{U}$ instead of $\mathbf{Log}\boldsymbol{U}$ (Model 4) is fully coherent with what one would anticipate based on the wage curve literature. Moreover, Model 3 is omitting important spatial effects, as the correlation between actual and predicted wages is lower.

The outcome of instrumenting $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{U}$ as well as the endogenous spatial lag (Model 5) is that the model's fit is reduced and the coefficient on the variable of interest becomes too large, in absolute terms, compared to available evidence of the wage curve. Therefore, Model 4 remains the preferred specification.

### *3.6.2 The Urban Economics (UE) model*

**Table 3.2 Results from the rival UE model estimated in isolation**

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| *Estimation method* | **S2SLS / GMM** | **S2SLS / GMM** | **Iterated [a] S2SLS / GMM** | **Iterated [a] S2SLS / GMM** |
| *Treatment of* (Log$E$-$\rho(\mathbf{I}_T \otimes \mathbf{W})$Log$E$) | **Exogenous** | **Exogenous** | **Exogenous** | **Endogenous** |
| In-commuting Labour Efficiency Endog. Spatial Lag ($\rho$) | | | **0.1049** | **0.1653** |
| (t-stat) | | | (7.79)*** | (6.33)*** |
| Local Employment Density (Log$E$-$\rho(\mathbf{I}_T \otimes \mathbf{W})$Log$E$) $(\gamma - 1)$ | **0.0348** | **0.0199** | **0.0220** | **0.1025** |
| (t-stat) | (7.87)*** | (5.28)*** | (6.13)*** | (4.86)*** |
| Local Tech. Knowledge Base Log$T$ ($a_2$) | | **0.0507** | **0.0307** | **-0.0270** |
| (t-stat) | | (10.79)*** | (6.60)*** | (-1.61)** |
| Local Unskilled Workforce Log$S$ ($a_3$) | | **-0.1774** | **-0.1781** | **-0.3234** |
| (t-stat) | | (-6.94)*** | (-7.70)*** | (-6.54)*** |
| Constant ($a_0$) | 5.9408 | 6.9562 | 6.1834 | 5.5344 |
| (t-stat) | (208.78)*** | (82.88)*** | (55.70)*** | (22.40) |
| *Error process* | | | | |
| λ [b] | **0.6898*** ** | **0.5354*** ** | **0.5303*** ** | **0.4519*** ** |
| $\sigma_v^2$ | 0.0039 | 0.0059 | 0.0051 | 0.0056 |
| $\sigma_1^2 = \sigma_v^2 + T\sigma_u^2$ | 0.0973 | 0.0692 | 0.0611 | 0.1561 |
| CORR [c] | **0.4320** | **0.6270** | **0.7109** | **0.6132** |
| No. areas | 408 | 408 | 408 | 408 |
| No. in-sample years (1999-2009) | 11 | 11 | 11 | 11 |

[a] Iteration is to satisfy the constraint involving $\rho$.
[b] Significance of the spatial autoregressive parameter is based on bootstrap inference (see Ch. 2).
[c] Correlation between observed and fitted values of **Log*Wage***.
* is significance at 10% level, ** is significance at 5% level, *** is significance at 1% level.

Table 3.2 summarises the outcome of estimating the rival (non-nested) UE model. The results in column 1 correspond to a basic UE specification, without controlling for local and in-commuting labour efficiency. The coefficient on **Log*E*** (here $\rho$ is constrained to zero) is strongly significant and its magnitude of 0.0348 is at the lower bound of the density elasticities typically found in the agglomeration literature, ranging 3-8% (Rosenthal and Strange, 2004); it means that doubling the number of employees per square kilometre increases local wages by 2.4% ($2^{0.0348}$-1=0.0244).

Allowing for local labour efficiency variations (column 2) reduces the size of the coefficient on $\mathbf{Log}\boldsymbol{E}$ by around one third, an indication that labour efficiency is not orthogonal to employment density, although evidence remains of a very significant employment density effect.

Commuting, as embodied in the endogenous spatial lag (column 3), also emerges as a significant determinant of local pay ( $\hat{\rho}$ =0.1049 with *t*-ratio=7.79), suggesting that wages are higher in areas with in-commuting flows of skilled and qualified workers (in line with the interpretation of the endogenous spatial lag as derived in section 3.3.3). Despite relevant labour efficiency spillovers, the efficiency level of the resident workforce is still a significant explanatory factor, with both $\mathbf{Log}\boldsymbol{T}$ and $\mathbf{Log}\boldsymbol{S}$ carrying the expected sign and remaining statistically significant; in particular, doubling a region's relative specialisation in computing and R&D activities ($\boldsymbol{T}$) raises wages by 2.15% ($2^{0.0307}$-1=0.0215), while a fall by a half in the proportion of working-age population without qualifications ( $\boldsymbol{S}$ ) produces an increase in wages of 13.14% ($2^{0.1781}$-1=0.1314). A separate source of higher wages is represented by increasing returns to employment density, as the estimate of $(\gamma - 1)$ also remains significantly above zero.

Model 3 gives a density elasticity of 0.022. Looking at similar studies of urban agglomeration, also using small-area data but in a cross-sectional rather than panel context, this is somewhat between the wage premium of around 0.015 for Great Britain (Fingleton, 2003; 2006) and that of 0.03-0.04 for France (Barde, 2010).

Instrumenting $(\mathbf{Log}\boldsymbol{E}\text{-}\rho(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{E})$ as well as the endogenous spatial lag (column 4) reduces the fit of the model and produces distorted estimates, with the coefficient on $(\mathbf{Log}\boldsymbol{E}\text{-}\rho(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{E})$ becoming too large compared to theoretical/ empirical expectations and that on $\mathbf{Log}\boldsymbol{T}$ turning negative. Therefore, Model 3 remains the preferred specification as it gives more plausible estimates.

The positive and significant estimate of $\lambda$ points to positive spatial residuals dependence due to common shocks or omitted (positively) spatially autocorrelated explanatory variables. Moreover, with a correlation between actual and predicted wages equal to 71%, the UE model (column 3) has the same level of fit as the wage curve (72%, see Table 3.1, column 4), which is an informal indication that neither model is quantitatively superior to the other. I carry out a more formal analysis subsequently.

### *3.6.3 The New Economic Geography (NEG) model*

**Table 3.3 Results from the rival NEG model estimated in isolation**

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| *Estimation method* | **S2SLS / GMM** | **S2SLS / GMM** | **Iterated [a] S2SLS / GMM** | **Iterated [a] S2SLS / GMM** |
| *Treatment of* $(\text{Log}MP\text{-}\rho(\mathbf{I}_T \otimes \mathbf{W})\text{Log}MP)$ | **Exogenous** | **Exogenous** | **Exogenous** | **Endogenous** |
| In-commuting Labour Efficiency | | | | |
| Endog. Spatial Lag ($\rho$) | | | **0.0876** | **0.0876** |
| (t-stat) | | | (6.92)*** | (6.93)*** |
| Local Market Potential | | | | |
| $(\text{Log}MP\text{-}\rho(\mathbf{I}_T \otimes \mathbf{W})\text{Log}MP)$ $(1/\varepsilon)$ | **0.2137** | **0.1225** | **0.1214** | **0.1422** |
| (t-stat) | (6.85)*** | (6.04)*** | (5.83)*** | (5.33)*** |
| Local Tech. Knowledge Base | | | | |
| Log$T$ ($a_2$) | | **0.0507** | **0.0396** | **0.0386** |
| (t-stat) | | (12.32)*** | (10.01)*** | (9.63)*** |
| Local Unskilled Workforce | | | | |
| Log$S$ ($a_3$) | | **-0.1343** | **-0.1207** | **-0.1176** |
| (t-stat) | | (-5.53)*** | (-5.49)*** | (-5.32)*** |
| Constant ($a_0$) | 4.4933 | 5.9806 | 5.4146 | 5.2518 |
| (t-stat) | (18.76)*** | (31.93)*** | (28.66)*** | (23.09)*** |
| *Error process* | | | | |
| $\lambda$ [b] | **0.7185*** ** | **0.5825*** ** | **0.5849*** ** | **0.5953*** ** |
| $\sigma_v^2$ | 0.0036 | 0.0052 | 0.0047 | 0.0046 |
| $\sigma_1^2 = \sigma_v^2 + T\sigma_u^2$ | 0.0830 | 0.0617 | 0.0569 | 0.0569 |
| CORR [c] | **0.4980** | **0.6415** | **0.6982** | **0.6935** |
| No. areas | 408 | 408 | 408 | 408 |
| No. in-sample years (1999-2009) | 11 | 11 | 11 | 11 |

[a] Iteration is to satisfy the constraint involving $\rho$.
[b] Significance of the spatial autoregressive parameter is based on bootstrap inference (see Ch. 2).
[c] Correlation between observed and fitted values of **Log*Wage***.
* is significance at 10% level, ** is significance at 5% level, *** is significance at 1% level.

Table 3.3 summarises the outcome of estimating the rival (non-nested) NEG model. The coefficient on market potential from a basic NEG specification is correctly (positively) signed and strongly significant (column 1) but half the size once local labour efficiency as proxied by **Log*T*** and **Log*S*** are controlled for (column 2).

Labour efficiency spillovers are also important (column 3), as can be inferred from a spatially autoregressive coefficient which is positively signed and statistically significant ( $\hat{\rho}$ =0.0876, *t*-ratio=6.92).

According to Model 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 S2SLS / GMM procedure but structurally estimating a short-run NEG wage equation. Fingleton (2008b) 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 (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, 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 (2008b) 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[26]. In the GB context, the study of Britain's UALADs by Fingleton (2006) gives a value of 0.15, which is very close to that found here; his essay 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 and in-commuting labour efficiency variations are similar.

Instrumenting $(\mathbf{Log}\boldsymbol{MP}\text{-}\rho(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{MP})$ as well as the endogenous spatial lag (Model 4) leaves estimates broadly unchanged, however I take Model 3 as the preferred specification for comparability with the wage curve and the UE model where the variable of interest – exogenised through a temporal lag transformation – is treated as orthogonal to errors.

[26] 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, Garretsen and Schramm, 2004 for Germany, and Niebuhr, 2006 for 158 European regions).

## 3.7 Models comparison and selection: 'Inclusive Regressions' approach

One important issue is that the wage curve and its rival 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 these non-nested competing hypotheses.

### *3.7.1. Asymptotic results from Artificial Nesting Models*

To shed light on which model might be the preferred specification from an econometric perspective, I initially adopt an 'Inclusive Regression' approach (Davidson and MacKinnon, 1993; Hendry, 1995). Thus, after estimating the models individually, I combine the wage curve and either the UE or NEG theory in a single empirical specification of which each model is a special case; this composite data generating process (DGP) is referred to in the literature as an Artificial Nesting Model (ANM). The problem amounts to testing whether there is a significant loss of information in reducing the ANM to either the wage curve or UE/ NEG by restricting either of these effects to zero.

I find that the wage curve is neither dominated by UE (Table 3.4) or NEG (Table 3.5), as unemployment retains its predictive power within the ANM, nor dominates them, as both employment density (Table 3.4) and market potential (Table 3.5) are also statistically significant. Therefore, the evidence from fitting an ANM is that neither rival encompasses the DGP - or, since the DGP nests both of the rivals, neither rival encompasses the other, in the sense that the predictive value of one cannot explain the results of the other - which means that both unemployment and either within- or between-region economic geography should be used to predict wages.

So far I have seen that both unemployment and either employment density or market potential should enter the wage equation. This is the case irrespective of whether local unemployment or unemployment within commuting distance is used, however there are two reasons for preferring an empirical specification with the latter variable. First, the coefficient on $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{U}$ is more in line with the classic elasticity of -0.10, as in Table 3.1. Second, omitting spatial inderdependence from unemployment variations when it clearly plays a role would seem to 'stack the odds' against the wage curve; I am explaining wages by workplace, so that what is important is unemployment within commuting distance, not unemployment of the place of residence (claimant counts data are residence-based). Employment data are workplace-based so this rationale is not relevant for employment density, while market potential is spatially lagged by definition / construction.

Moreover I see that the coefficients on $\mathbf{Log}\boldsymbol{T}$ and $\mathbf{Log}\boldsymbol{S}$ are appropriately signed and statistically significant, a result which endorses an extended empirical specification as set out in section 3.3.3 that controls for labour efficiency. The spatially lagged dependent variable is also highly relevant, and improves the fit of the model 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 market potential within an ANM; in a later study seeking to explain individual-level, home-based wages from the British Household Panel Survey database, Fingleton and Longhi (2013) estimate an ANM which combines per-district, within-commuting-distance employment density and market potential – also controlling for local unemployment and a set of individual-level attributes (e.g. age, marriage, children) - and find that, having taken spatial effects into account, market potential is not a factor affecting pay.

### *3.7.2. Bootstrap results from 'Inclusive Regressions'*

Further insights can be gained using a bootstrap *t*-ratio reference distribution.

In Table 3.6, I present bootstrap results for the Artificial Nesting Models in the last columns of Tables 3.4 and 3.5 (both incorporating an endogenous spatial lag, use unemployment rate within commuting distance as wage curve hypothesis, and assume exogeneity for the rival regressors). Using cross-sectional notation:

1. I start by generating simulated wage data via $\hat{\rho}^{ANM}, \hat{\beta}^{ANM}, \hat{\lambda}^{ANM}, \hat{\sigma}_u^{2ANM}, \hat{\sigma}_v^{2ANM}$ (parameter estimates from the ANM specifications, respectively for wage curve *vs* UE and wage curve *vs* NEG, as reported in the last columns of Tables 3.4 and 3.5)

and with simulated values for error components $u^{ANM}$ and $v^{ANM}$ obtained by randomly drawing from a Normal distribution with zero mean and estimated variances $\hat{\sigma}_u^{2ANM}$ and $\hat{\sigma}_v^{2ANM}$ .

2. I then fit the ANM specifications to these simulated wage data by the usual S2SLS / GMM estimation procedure.

3. Steps 1-2 are repeated $B$ = 999 times, and the bootstrap $P$-values associated with $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{U}$ and $(\mathbf{Log}\boldsymbol{E}\text{-}\rho(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{E})$ , or $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{U}$ and $(\mathbf{Log}\boldsymbol{MP}\text{-}\rho(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{MP})$ , are calculated as the fraction of times that the simulated $t$-ratios are more extreme than the estimated $t$-ratios as reported in the last columns of Tables 3.4 and 3.5 (i.e. more negative for $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Log}\boldsymbol{U}$ , or more positive for employment density and market potential).

With regard to wage curve *vs* UE, bootstrap results confirm that both spatially-lagged unemployment and employment density are statistically significant under the ANM, meaning that excluding either of these variables would significantly reduce the model's fit. With regard to wage curve *vs* NEG, bootstrap evidence points to the superiority of the wage curve; the simulated $P$-value on market potential (0.61) indicates that this has no predictive power within the ANM, while the bootstrap $P$-value on spatially-lagged unemployment (0.14) is much smaller and closer to a conventional 10% level required for statistical significance.

With regard to which of the rival UE and NEG theories is more challenging for the wage curve, bootstrap results suggest that UE is the stronger competing paradigm. This is in line with consensus of the empirical geographical economics literature that, at lower levels of spatial aggregation, market potential has weaker explanatory force while employment density is more relevant (Brülhart and Mathys, 2008; Brakman, Garretsen and Van Marrewijk, 2009b). For example, in the context of Britain's UALADs, Fingleton's (2006) ANM results show that UE is superior to NEG whereas, in the case of EU NUTS2 regions (which are somewhat larger), Fingleton (2007) finds that both are acceptable.

Bootstrap analysis allows direct comparability between the ANM results in this chapter and the bootstrap $J$-test results in the next chapter.

**Table 3.4 Results from 'inclusive regressions' nesting Wage Curve and UE models**

| | 1 |
|---|---|
| *Estimation method* | **Iterated [a] S2SLS / GMM** |
| *Treatment of* $(\text{Log}E\text{-}\rho(\mathbf{I}_T \otimes \mathbf{W})\text{Log}E)$ | **Exogenous** |
| Endog. Spatial Lag | **0.1221** |
| (t-stat) | (9.09)*** |
| Unemp. within Commuting Distance | |
| $(\mathbf{I}_T \otimes \mathbf{W})\text{Log}U$ | **-0.1256** |
| (t-stat) | (-6.55)*** |
| Local Employment Density | |
| $(\text{Log}E\text{-}\rho(\mathbf{I}_T \otimes \mathbf{W})\text{Log}E)$ | **0.0195** |
| (t-stat) | (5.20)*** |
| Local Tech. Knowledge Base | |
| Log$T$ | **0.0265** |
| (t-stat) | (5.66)*** |
| Local Unskilled Workforce | |
| Log$S$ | **-0.1557** |
| (t-stat) | (-6.48)*** |
| Constant | 6.0710 |
| (t-stat) | (54.47)*** |
| *Error process* | |
| λ [b] | **0.6159*** ** |
| $\sigma_v^2$ | 0.0036 |
| $\sigma_1^2 = \sigma_v^2 + T\sigma_u^2$ | 0.0619 |
| CORR [c] | **0.7507** |
| No. areas | 408 |
| No. in-sample years (1999-2009) | 11 |

[a] Iteration is to satisfy the constraint involving $\rho$.
[b] Significance of the spatial autoregressive parameter is based on bootstrap inference (see Ch. 2).
[c] Correlation between observed and fitted values of **Log*Wage***.
* is significance at 10% level, ** is significance at 5% level, *** is significance at 1% level.

**Table 3.5 Results from 'inclusive regressions' nesting Wage Curve and NEG models**

| | 1 |
|---|---|
| *Estimation method* | **Iterated [a] S2SLS / GMM** |
| *Treatment of* (Log$MP$-$\rho$($\mathbf{I}_T \otimes \mathbf{W}$)Log$MP$) | **Exogenous** |
| Endog. Spatial Lag | **0.1113** |
| (t-stat) | (8.85)*** |
| Unemp. within Commuting Distance | |
| ($\mathbf{I}_T \otimes \mathbf{W}$)Log$U$ | **-0.1221** |
| (t-stat) | (-6.02)*** |
| Local Market Potential | |
| (Log$MP$-$\rho$($\mathbf{I}_T \otimes \mathbf{W}$)Log$MP$) | **0.1118** |
| (t-stat) | (4.38)*** |
| Local Tech. Knowledge Base | |
| Log$T$ | **0.0354** |
| (t-stat) | (9.04)*** |
| Local Unskilled Workforce | |
| Log$S$ | **-0.1086** |
| (t-stat) | (-4.91)*** |
| Constant | 5.3682 |
| (t-stat) | (24.78)*** |
| *Error process* | |
| $\lambda$ [b] | **0.6813*** ** |
| $\sigma_v^2$ | 0.0034 |
| $\sigma_1^2 = \sigma_v^2 + T\sigma_u^2$ | 0.0550 |
| CORR [c] | **0.7443** |
| No. areas | 408 |
| No. in-sample years (1999-2009) | 11 |

[a] Iteration is to satisfy the constraint involving $\rho$.
[b] Significance of the spatial autoregressive parameter is based on bootstrap inference (see Ch. 2).
[c] Correlation between observed and fitted values of **Log*Wage***.
* is significance at 10% level, ** is significance at 5% level, *** is significance at 1% level.

## Table 3.6 Bootstrap *P*-values for key variables in 'inclusive regressions'

| | Wage Curve vs UE | | Wage Curve vs NEG | |
|---|---|---|---|---|
| | **Table 3.4, Model 4** | **Bootstrap** | **Table 3.5, Model 4** | **Bootstrap** |
| Unemp. within Commuting Distance | | | | |
| $(\mathbf{I}_T \otimes \mathbf{W})\text{Log}U$ | -0.1256 | -0.1256 | -0.1221 | -0.1221 |
| **(*P*-value)** | **(0.000)** | **(0.003)** | **(0.000)** | **(0.138)** |
| Local Employment Density | | | | |
| $(\text{Log}E\text{-}\rho(\mathbf{I}_T \otimes \mathbf{W})\text{Log}E)$ | 0.0195 | 0.0195 | | |
| **(*P*-value)** | **(0.000)** | **(0.011)** | | |
| Local Market Potential | | | | |
| $(\text{Log}MP\text{-}\rho(\mathbf{I}_T \otimes \mathbf{W})\text{Log}MP)$ | | | 0.1118 | 0.1118 |
| **(*P*-value)** | | | **(0.000)** | **(0.608)** |
| **Inference about the relative explanatory performance of the two rival hypotheses when confronted directly** | **Significant for both models** | **Significant for both models** | **Significant for both models** | **Dominance of Wage Curve** |

## 3.8 Conclusions

Looking at Britain's 408 local authorities over the period 1999-2009, the present chapter has explored the predicting performance of the wage curve relative to the Urban Economics (UE) and New Economic Geography (NEG) considered in turn, in order to establish whether the wage curve truly represents an empirical reality.

I estimated an inclusive wage equation that nests both, to see which model (if any) encompasses the other. When using asymptotic *P*-values, I find that unemployment is not nullified by either UE or NEG, 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 either employment density or market potential also entails a significant loss of information. Meanwhile, when using bootstrap *P*-values, the wage curve emerges as superior to NEG, while inference about its performance relative to UE remains unchanged.

These results point to the conclusion that, at least in such small-scale and short-distance spatial context, although the wage curve holds it should not be taken as an absolute principle governing the spatial distribution of economic development; there are in fact other strands of regional economics – most notably Urban Economics - which are able to account for local wage rate variance, and these are firmly grounded in economic theory as well as being validated empirically.

# Chapter 4

# Bootstrap *J*-Test for Panel Data Models with Spatially Dependent Error Components, a Spatial Lag and Additional Endogenous Variables

## 4.1 Introduction

In recent years there has been an increasing interest in the spatial panel literature, however hypothesis testing procedures in spatial panel econometrics are somewhat lacking. To fill this gap, this essay extends existing work on the *J*-test for spatial fixed-effects models to spatial models (SARAR-RE models) incorporating random effects (also subjected, like the disturbances, to a spatially autoregressive process) as well as a spatial lag of the dependent variable and additional, potentially endogenous regressors.

The *J*-test is a well-established technique for testing a null model (*H0*) against one or more non-nested alternatives (*H1*); there is a large literature concerning a non-spatial cross-sectional framework (Davidson and MacKinnon, 1981; MacKinnon et al., 1983; Dastoor, 1983; Godfrey, 1983; Godfrey and Pesaran, 1983; Delgado and Stengos, 1994). Among all non-nested hypotheses tests, the *J*-test is the most widely used (MacAleer, 1995), for its intuitive appeal and because it compares well with similar tests as indicated by Monte Carlo simulations presented in Davidson and MacKinnon (1982). The *J*-test is strongly grounded in the non-nested test literature being based on the encompassing principle of Mizon and Richard (1986), which addresses the question of whether a maintained model can explain the features of competing alternatives. The test looks at whether fitted values based on the rival model add significantly to the explanatory capability of the null model, the rationale being that if the *H0* model is correct then none of the variation in the regressand would be captured by the augmenting *H1* predictions.

In the non-spatial cross-sectional context, a number of studies have demonstrated that there is a finite-sample problem with the asymptotic use of the *J*-test, namely that the finite-sample null reference distribution of the *J*-statistic lies to the right of its asymptotic distribution, which is N(0,1). This means that, in small samples, the Standard Normal is not valid as a reference distribution for the *J*-statistic, as it leads to size distortion (an inflated Type I error rate compared to nominal levels) and excessive rejection of the null model. Monte Carlo results in Fan and Li (1995), Godfrey (1998) and Davidson and MacKinnon (2002) have all suggested that the bootstrap can improve the finite-sample properties of the test and allow the procedure to be usefully applied also in small samples.

I implement a version of the *J*-test for SARAR-RE models building on work by Kelejian (2008), Burridge and Fingleton (2010) and Kelejian and Piras (2011), who initially generalised Davidson and MacKinnon's (1981) test to spatial cross-sectional data, and particularly Kelejian and Piras (2015) who present a *J*-test procedure for panel data models that include an endogenous spatial lag and additional endogenous variables in a fixed-effect setting. The spatial *J*-test in this paper is also for panel data models but in cases where the specification of the null, and alternative, hypotheses account for spatial heterogeneity via random effects. Random effects (RE) have various benefits over fixed effects (FE) and in many cases as the present one are the appropriate and preferred option, for example because they allow identification of time-constant covariates and because they capture both within- and between-sample variation instead of solely relying on time-series information; it is important to emphasise the RE approach as a distinct option rather than be treated as somehow subsidiary or unnecessary given the existence of a FE specification. A random-effects approach thus represents a useful addition to the spatial econometric literature on non-nested hypotheses testing when rival models are fitted to panel data, although it is not the focus of the analysis nor the main justification for this study.

The emphasis of this paper is rather on the parametric bootstrap for SARAR-RE models developed in order to control rejection frequencies in small samples (i.e. to provide a better approximation of the (unknown) "true" cumulative distribution function of the *J*-statistic under the null). The reason for this, in addition to the finite-sample problem, is that SARAR-RE estimation has not been considered in the literature on the $J$-test to date, therefore it is not obvious that asymptotic critical values will guarantee correct inference. In a spatial cross-sectional scenario, the application of bootstrapping to the $J$-test is advocated by Burridge and Fingleton (2010), while Kelejian and Piras (2011) focus on asymptotics but accept that in practice small-sample inferences based on a bootstrap approach can yield more reliable results. In a spatial panel-data (fixed-effects) scenario, Kelejian and Piras (2015) find that Type I errors are very reasonable in moderately-sized samples; I thus report empirical size and power comparisons of small-to-medium sample approximations to the “true” distribution of the $J$-statistic based on the bootstrap method and based on asymptotic theory, proving that the bootstrap distribution is correct and strengthening the case for bootstrapping (which itself is quite conventional and has a reasonable pedigree with very small samples).

With regards to estimation, I fit the models via the Spatial Two-Stage Least Squares / Generalised Method of Moments (S2SLS / GMM) of Kapoor et al. (2007) and Fingleton (2008a), choosing this method over alternative strategies for random-effects models such as the four-step procedure based on within as well as between residuals suggested by Piras (2013). The motivation is that, with a very small time dimension ($T$=11) compared to the cross-sectional dimension ($N$=408), the bulk of the variation is between- rather than within-sample, so it is reasonable to think that the within transformation would not be particularly advantageous in this case.

The bootstrap $J$-test is illustrated using non-nested wage equations - namely the spatial wage curve, which explains local pay variation with reference to unemployment within commuting distance, and either the NEG or UE theory, respectively referring to market potential and employment density. In line with bootstrap ANM results in the previous chapter, bootstrap $J$-test results point to the wage curve being the most appropriate model of wage determination for Britain’s UALADs when tested against NEG but not when tested against UE.

In Section 4.2 I describe the history of the problem. In Section 4.3 I specify the $J$-test procedure for SARAR-RE models while the empirical set-up is outlined in Section 4.4. Section 4.5 describes the parametric bootstrap design for finite-sample inference with SARAR-RE models, with results discussed in Sections 4.6 and 4.7 for the empirical application and for the Monte Carlo experiments respectively. Section 4.8 summarises and concludes.

### 4.2 History of the problem

In this section I focus first on non-spatial models, setting the context for the history of the problem given spatial data.

Fan and Li (1995) is one of the non-spatial cross-sectional contributions to establish that the use of asymptotic critical values for *J* can cause misleading outcomes, as the *J*-test has significance levels in small samples which are often considerably greater than the nominal size. They adopt a residuals resampling approach and demonstrate that the bootstrap provides a better measure of the finite-sample distribution for the *J*-statistic than its asymptotic Standard Normal approximation. Moreover, they draw attention to the fact that the *J*-test suffers from size distortion under conditions of near orthogonality between the rival hypotheses, and show that the bootstrap can cure this problem also. Their study, however, is somewhat limited in that the data processes in the Monte Carlo replicates all have normally and identically distributed disturbances.

Godfrey (1998) also discusses how the use of N(0,1) critical values can severely overreject a true null model and, by means of residuals resampling, shows that bootstrapping can substantially reduce the problem of size inflation. Further, he proves that the bootstrap is robust to error distribution assumptions, with observed rejection frequencies closely agreeing with the nominal 5% probability in the presence of either normal or non-normal shocks.

The findings of Fan and Li (1995) and Godfrey (1998) are reinforced by Davidson and MacKinnon (2002). In the non-spatial context, they find that in extreme cases, e.g. for sample sizes of fifty, an asymptotic *J*-test at the 5% significance level can reject a true null hypothesis more than 80% of the time. Their Monte Carlo replications also indicate that the bootstrap *J*-test works well in small samples compared with the ordinary test, regardless of whether the assumptions of regressors exogeneity and shock normality hold or not.

With reference to GMM-type estimation of spatial cross-sectional models, Burridge and Fingleton (2010) determine that in small-to-medium samples the asymptotic $J$ -test of Kelejian (2008) can be too liberal in various parts of the parameter space, and that in most cases the empirical significance levels can be corrected by the use of the bootstrap to construct a valid reference distribution. Their experiments relate to the case of a single alternative model and a single non-constant regressor, with either different weights matrices or different explanatory factors but not both. In a similar exercise, Burridge (2012) illustrates how sensitive the test's properties are to spatial structure, and finds that significance levels are not greatly influenced by the form of the weights.

This strand of research thus indicates that statistical inference should be based on *P*-values computed using a bootstrap distribution which is constructed by simulation, rather than from a reference distribution which is suggested by large-sample theory, as this tool can yield more accurate outcomes than traditional approaches.

In a spatial fixed-effects framework, the results in Kelejian and Piras (2015) show that the $J$ -test has good power and empirical size reasonably close to the theoretical 5% level for moderately-sized samples, but the true Type I error rates are not always close to the nominal level when the sample size is very small. Therefore, I use Monte Carlo trials in the context of S2SLS / GMM estimation to show the finite-sample properties of the bootstrap $J$ -test and to establish whether asymptotics would be adequate in small-to-medium samples or a bootstrap distribution is actually needed.

## 4.3 The *J*-test procedure for SARAR-RE models

### *4.3.1 Null and alternative hypotheses*

Under the null hypothesis, SARAR-RE model *H0* is true

$$\begin{aligned} \mathbf{Y} &= \iota_0 \boldsymbol{\iota}_{TN} + \rho_0 (\mathbf{I}_T \otimes \mathbf{W})\mathbf{Y} + \alpha_0 \mathbf{Z}_0 + \mathbf{X}_0 \boldsymbol{\beta}_0 + \mathbf{e}_0 \\ \mathbf{e}_0 &= \lambda_0 (\mathbf{I}_T \otimes \mathbf{M})\mathbf{e}_0 + \boldsymbol{\xi}_0 \\ \boldsymbol{\xi}_0 &= (\boldsymbol{\iota}_T \otimes \mathbf{u}_0) + \mathbf{v}_0 \end{aligned} \tag{4.1a}$$

where $\boldsymbol{\iota}_{TN}$ is a *TN* x 1 vector of ones (i.e. a constant term), $\mathbf{Y}$ is a *TN* x 1 vector of observations on the dependent variable (i.e. local weekly wage rates); $\mathbf{Z}_0$ is a *TN* x 1 vector of observations on the null model's endogenous explanatory variable (spatial lag of the unemployment rate); $\mathbf{X}_0$ is a *TN* x $k_0$ matrix of observations on the null model's $k_0$ exogenous regressors (local technical knowledge and local (lack of) educational attainment); $\mathbf{W}$ and $\mathbf{M}$ are the *N* x *N* non-stochastic pre-defined matrices of exogenous spatial weights; $\mathbf{I}_T$ is a *T* x *T* diagonal matrix with ones on the main diagonal and zeros elsewhere, and $\boldsymbol{\iota}_T$ is a *T* x 1 vector of ones. The parameters to be estimated are the intercept $\iota_0$, the slope coefficient $\alpha_0$, the slope coefficients in the $k_0$ x 1 vector $\boldsymbol{\beta}_0$, the spatial autoregressive parameters $\rho_0$ and $\lambda_0$, and the error variances $\sigma^2_{0u}$ and $\sigma^2_{0v}$.

Moreover, $\mathbf{e}_0$ is a *TN* x 1 vector of spatially dependent error terms, and $\boldsymbol{\xi}_0$ is the usual *TN* x 1 vector of innovations which combines a permanent, i.e. time-constant, error component $\mathbf{u}_0 \sim iid(0,\sigma^2_{0u})$ and a transient, i.e. time-varying, error component $\mathbf{v}_0 \sim iid(0,\sigma^2_{0v})$, respectively a random-effects vector picking up unobserved or unmeasured (time-invariant) interregional heterogeneity and a disturbances or shocks vector. The covariance matrix for $\boldsymbol{\xi}_0$ is $\boldsymbol{\Omega}_{0\zeta} = \sigma^2_{0u}(\boldsymbol{J}_T \otimes \boldsymbol{I}_N) + \sigma^2_{0v}\boldsymbol{I}_{TN}$, where $\boldsymbol{J}_T$ is a *T* x *T* matrix of ones. Both $\mathbf{u}_0$ and $\mathbf{v}_0$ are subject to the 'same' spatial autoregressive process (see eq. 2.10), as evident from

$$\begin{aligned}\mathbf{e}_0 &= (\mathbf{I}_{TN} - \lambda_0(\mathbf{I}_T \otimes \mathbf{M}))^{-1}\boldsymbol{\xi}_0 = \\ &= (\mathbf{I}_{TN} - \lambda_0(\mathbf{I}_T \otimes \mathbf{M}))^{-1}(\boldsymbol{\iota}_T \otimes \mathbf{u}_0) + (\mathbf{I}_{TN} - \lambda_0(\mathbf{I}_T \otimes \mathbf{M}))^{-1}\mathbf{v}_0 = \\ &= (\boldsymbol{\iota}_T \otimes (\mathbf{I}_N - \lambda_0\mathbf{M})^{-1})\mathbf{u}_0 + (\mathbf{I}_T \otimes (\mathbf{I}_N - \lambda_0\mathbf{M})^{-1})\mathbf{v}_0\end{aligned} \tag{4.1b}$$

Under the alternative, the data are generated by a similar structure, giving SARAR-RE model *H1*

$$\begin{aligned}\mathbf{Y} &= \iota_1\boldsymbol{\iota}_{TN} + \rho_1(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Y} + \alpha_1\mathbf{Z}_1 + \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{e}_1 \\ \mathbf{e}_1 &= \lambda_1(\mathbf{I}_T \otimes \mathbf{M})\mathbf{e}_1 + \boldsymbol{\xi}_1 \\ \boldsymbol{\xi}_1 &= (\boldsymbol{\iota}_T \otimes \mathbf{u}_1) + \mathbf{v}_1\end{aligned} \tag{4.2}$$

where $\mathbf{Z}_1$ is a $TN$ x 1 vector of observations on the alternative model's endogenous explanatory variable (market potential or employment density), and $\mathbf{X}_1$ is a $TN$ x $k_1$ matrix of observations on the alternative model's $k_1$ exogenous regressors (local technical knowledge and local (lack of) educational attainment). The spatial autoregressive processes involving $\mathbf{Y}$ and $\mathbf{e}_1$, and the exogenous weighting matrices $\mathbf{W}$ and $\mathbf{M}$ which govern them, are identical to those in model *H0*. I can write

$$\begin{aligned} H0: \mathbf{Y} &= \iota_0 \boldsymbol{\iota}_{TN} + \rho_0(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Y} + \alpha_0 \mathbf{Z}_0 + \mathbf{X}_0 \boldsymbol{\beta}_0 + \mathbf{e}_0 = \mathbf{R}_0 \boldsymbol{\gamma}_0 + \mathbf{e}_0 \\ H1: \mathbf{Y} &= \iota_1 \boldsymbol{\iota}_{TN} + \rho_1(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Y} + \alpha_1 \mathbf{Z}_1 + \mathbf{X}_1 \boldsymbol{\beta}_1 + \mathbf{e}_1 = \mathbf{R}_1 \boldsymbol{\gamma}_1 + \mathbf{e}_1 \end{aligned} \tag{4.3}$$

in which $\boldsymbol{\gamma}_0 = \left[\iota_0, \rho_0, \alpha_0, \boldsymbol{\beta}_0'\right]'$ and $\boldsymbol{\gamma}_1 = \left[\iota_1, \rho_1, \alpha_1, \boldsymbol{\beta}_1'\right]'$, and also $\mathbf{R}_0 = \left[\boldsymbol{\iota}_{TN}, (\mathbf{I}_T \otimes \mathbf{W})\mathbf{Y}, \mathbf{X}_0, \mathbf{Z}_0\right]$ and $\mathbf{R}_1 = \left[\boldsymbol{\iota}_{TN}, (\mathbf{I}_T \otimes \mathbf{W})\mathbf{Y}, \mathbf{X}_1, \mathbf{Z}_1\right]$. Here I have two exogenous covariates both of which, in this particular application (see section 4.4), are common to the null and alternative models, so that $\mathbf{X}_0 = [\mathbf{X}_{01}, \mathbf{X}_{02}]$ and $\mathbf{X}_1 = [\mathbf{X}_{11}, \mathbf{X}_{12}]$ are the same. More generally I have used different notation in eqs. (4.1) and (4.2) to signify that the exogenous covariates can be different.

### *4.3.2 Test specification*

I implement the test in four steps.

In *Step 1*, $\mathbf{e}_0$ and $\mathbf{e}_1$ are consistently estimated by the 2SLS / IV method. Following Kelejian and Piras (2011), I define matrices $\mathbf{L}_{0,r}$ for the null model, $\mathbf{L}_{1,r}$ for the alternative model, and $\mathbf{L}_{01,r}$ for the hybrid model that combines both, that is

$$\mathbf{L}_{0,r} = \left[\boldsymbol{\iota}_{TN} \vdots \mathbf{X}_{01} \vdots \mathbf{X}_{02} \vdots (\mathbf{I}_T \otimes \mathbf{W})\mathbf{X}_{01} \vdots \ldots \vdots (\mathbf{I}_T \otimes \mathbf{W}^r)\mathbf{X}_{01} \vdots (\mathbf{I}_T \otimes \mathbf{W})\mathbf{X}_{02} \vdots \ldots \vdots (\mathbf{I}_T \otimes \mathbf{W}^r)\mathbf{X}_{02}\right]$$

$$\mathbf{L}_{1,r} = \left[\boldsymbol{\iota}_{TN} \vdots \mathbf{X}_{11} \vdots \mathbf{X}_{12} \vdots (\mathbf{I}_T \otimes \mathbf{W})\mathbf{X}_{11} \vdots \ldots \vdots (\mathbf{I}_T \otimes \mathbf{W}^r)\mathbf{X}_{11} \vdots (\mathbf{I}_T \otimes \mathbf{W})\mathbf{X}_{12} \vdots \ldots \vdots (\mathbf{I}_T \otimes \mathbf{W}^r)\mathbf{X}_{12}\right]$$

$$\mathbf{L}_{01,r} = \left[\boldsymbol{\iota}_{TN} \vdots [\mathbf{X}_0 \vdots \mathbf{X}_1] \vdots (\mathbf{I}_T \otimes \mathbf{W})[\mathbf{X}_{01} \vdots \mathbf{X}_{02} \vdots \mathbf{X}_{11} \vdots \mathbf{X}_{12}] \vdots \ldots \vdots (\mathbf{I}_T \otimes \mathbf{W}^r)[\mathbf{X}_{01} \vdots \mathbf{X}_{02} \vdots \mathbf{X}_{11} \vdots \mathbf{X}_{12}]\right]_{LI}$$

where I use subscript *LI* to denote a spanning set of linearly independent columns, and allow spatial weights matrices $\mathbf{W}$ of any order up to some arbitrary small integer $r$ (Kelejian and Piras, 2011).

I then construct matrices of instruments $\mathbf{H}_{0,r} = [\mathbf{L}_{0,r} \vdots (\mathbf{I}_T \otimes \mathbf{M})\, \mathbf{L}_{0,r}]_{LI}$ , $\mathbf{H}_{1,r} = [\mathbf{L}_{1,r} \vdots (\mathbf{I}_T \otimes \mathbf{M})\, \mathbf{L}_{1,r}]_{LI}$ and $\mathbf{H}_{01,r} = [\mathbf{L}_{01,r} \vdots (\mathbf{I}_T \otimes \mathbf{M})\mathbf{L}_{01,r}]_{LI}$, and obtain projection matrices $\mathbf{P}_{0,r} = \mathbf{H}_{0,r}(\mathbf{H}'_{0,r}\mathbf{H}_{0,r})^{-1}\mathbf{H}'_{0,r}$ and $\mathbf{P}_{1,r} = \mathbf{H}_{1,r}(\mathbf{H}'_{1,r}\mathbf{H}_{1,r})^{-1}\mathbf{H}'_{1,r}$ which lead to IV estimators for *H0* and *H1* respectively

$$\hat{\boldsymbol{\gamma}}_{0,\mathrm{IV}} = \left[\mathbf{R}'_0\mathbf{P}_0\mathbf{R}_0\right]^{-1}\mathbf{R}'_0\mathbf{P}_0\mathbf{Y} \tag{4.4a}$$

$$\hat{\boldsymbol{\gamma}}_{1,\mathrm{IV}} = \left[\mathbf{R}'_1\mathbf{P}_1\mathbf{R}_1\right]^{-1}\mathbf{R}'_1\mathbf{P}_1\mathbf{Y} \tag{4.4b}$$

In *Step 2* the vectors of consistent residuals from IV estimation of the null and alternative models, defined as $\hat{\mathbf{e}}_0 = \mathbf{Y} - \mathbf{R}_0\hat{\boldsymbol{\gamma}}_{0,\mathrm{IV}}$ and $\hat{\mathbf{e}}_1 = \mathbf{Y} - \mathbf{R}_1\hat{\boldsymbol{\gamma}}_{1,\mathrm{IV}}$ , are used to estimate error process parameters $\lambda_0$, $\lambda_1$, $\sigma^2_{0u}$, $\sigma^2_{0\nu}$, $\sigma^2_{1u}$, and $\sigma^2_{1\nu}$ via the non-linear GM method of Kapoor et al. (2007).

In *Step 3*, I use $\hat{\lambda}_0$ and $\hat{\lambda}_1$ to carry out spatial Cochrane-Orcutt transformation

$$\begin{aligned}
\mathbf{Y}^*_0 &= (\mathbf{I}_{TN} - \hat{\lambda}_0(\mathbf{I}_T \otimes \mathbf{M}))\mathbf{Y} \\
\mathbf{R}^*_0 &= (\mathbf{I}_{TN} - \hat{\lambda}_0(\mathbf{I}_T \otimes \mathbf{M}))\mathbf{R}_0 \\
\mathbf{e}^*_0 &= (\mathbf{I}_{TN} - \hat{\lambda}_0(\mathbf{I}_T \otimes \mathbf{M}))\mathbf{e}_0
\end{aligned} \tag{4.5a}$$

$$\begin{aligned}
\mathbf{Y}^*_1 &= (\mathbf{I}_{TN} - \hat{\lambda}_1(\mathbf{I}_T \otimes \mathbf{M}))\mathbf{Y} \\
\mathbf{R}^*_1 &= (\mathbf{I}_{TN} - \hat{\lambda}_1(\mathbf{I}_T \otimes \mathbf{M}))\mathbf{R}_1 \\
\mathbf{e}^*_1 &= (\mathbf{I}_{TN} - \hat{\lambda}_1(\mathbf{I}_T \otimes \mathbf{M}))\mathbf{e}_1
\end{aligned} \tag{4.5b}$$

and construct the spatially lag-transformed regressions

$$\begin{gathered}
(\mathbf{I}_{TN} - \hat{\lambda}_0(\mathbf{I}_T \otimes \mathbf{M}))\mathbf{Y} = (\mathbf{I}_{TN} - \hat{\lambda}_0(\mathbf{I}_T \otimes \mathbf{M}))(\mathbf{R}_0\boldsymbol{\gamma}_0 + \mathbf{e}_0) \\
\mathbf{Y}^*_0 = \mathbf{R}^*_0\boldsymbol{\gamma}_0 + \mathbf{e}^*_0
\end{gathered} \tag{4.6a}$$

$$\begin{gathered}
(\mathbf{I}_{TN} - \hat{\lambda}_1(\mathbf{I}_T \otimes \mathbf{M}))\mathbf{Y} = (\mathbf{I}_{TN} - \hat{\lambda}_1(\mathbf{I}_T \otimes \mathbf{M}))(\mathbf{R}_1\boldsymbol{\gamma}_1 + \mathbf{e}_1) \\
\mathbf{Y}^*_1 = \mathbf{R}^*_1\boldsymbol{\gamma}_1 + \mathbf{e}^*_1
\end{gathered} \tag{4.6b}$$

Instruments sets $\mathbf{H}_{0,r}$ and $\mathbf{H}_{1,r}$ alongside estimated covariance matrices $\hat{\boldsymbol{\Omega}}_{0\zeta} = \hat{\sigma}^2_{0u}(\boldsymbol{J}_T \otimes \boldsymbol{I}_N) + \hat{\sigma}^2_{0\nu}\boldsymbol{I}_{TN}$ and $\hat{\boldsymbol{\Omega}}_{1\zeta} = \hat{\sigma}^2_{1u}(\boldsymbol{J}_T \otimes \boldsymbol{I}_N) + \hat{\sigma}^2_{1\nu}\boldsymbol{I}_{TN}$ give projection matrices $\mathbf{P}^*_0 = \left[\mathbf{H}_{0,r}(\mathbf{H}'_{0,r}\hat{\boldsymbol{\Omega}}_{0\xi}\mathbf{H}_{0,r})^{-1}\mathbf{H}'_{0,r}\right]$ and $\mathbf{P}^*_1 = \left[\mathbf{H}_{1,r}(\mathbf{H}'_{1,r}\hat{\boldsymbol{\Omega}}_{1\xi}\mathbf{H}_{1,r})^{-1}\mathbf{H}'_{1,r}\right]$ for the null and alternative models respectively. Given these, I obtain parameter estimates, fitted values and residuals vectors for the maintained hypothesis *H0*

$$\begin{aligned} \hat{\boldsymbol{\gamma}}_{0,S2SLS/GMM} &= \left[ \mathbf{R}_0^{*\prime} \mathbf{P}_0^* \mathbf{R}_0^* \right]^{-1} \mathbf{R}_0^{*\prime} \mathbf{P}_0^* \mathbf{Y}_0^* \\ \hat{\mathbf{Y}}_0^* &= \mathbf{R}_0^* \hat{\boldsymbol{\gamma}}_{0,2SLS/GMM} \\ \hat{\mathbf{e}}_0 &= \mathbf{Y} - \hat{\mathbf{Y}}_0^* \end{aligned} \tag{4.7a}$$

and the competing hypothesis *H1*

$$\begin{aligned} \hat{\boldsymbol{\gamma}}_{1,S2SLS/GMM} &= \left[ \mathbf{R}_1^{*\prime} \mathbf{P}_1^* \mathbf{R}_1^* \right]^{-1} \mathbf{R}_1^{*\prime} \mathbf{P}_1^* \mathbf{Y}_1^* \\ \hat{\mathbf{Y}}_1^* &= \mathbf{R}_1^* \hat{\boldsymbol{\gamma}}_{1,S2SLS/GMM} \\ \hat{\mathbf{e}}_1 &= \mathbf{Y} - \hat{\mathbf{Y}}_1^* \end{aligned} \tag{4.7b}$$

Kelejian and Piras (2011) suggest that, because the dependent and explanatory variables in the hybrid model are the outcome of a spatial Cochrane-Orcutt transformation involving $\hat{\lambda}_0$ , the *H1* predictions should be spatially lag-transformed in a similar way before being added to the maintained model. Their Monte Carlo simulations demonstrate that this version of the *J*-test, with the improved information set for the hybrid model, is more efficient than Kelejian's (2008) original procedure, and has more power to correctly distinguish between competing hypotheses (it correctly rejects the false null model more frequently). Following Kelejian and Piras (2011), I thus take the vector of fitted values $\hat{\mathbf{Y}}_1^*$ corresponding to the alternative model and apply the spatial C-O transformation using $\hat{\lambda}_0$ .

$$\hat{\mathbf{Y}}_1^*(\hat{\lambda}_0) = (\mathbf{I}_{TN} - \hat{\lambda}_0(\mathbf{I}_{\boldsymbol{T}} \otimes \mathbf{M}))\hat{\mathbf{Y}}_1^* \tag{4.8}$$

Then I augment the right-hand side of the null model in eq. (4.6a) with predictions $\hat{\mathbf{Y}}_1^*(\hat{\lambda}_0)$, which approximate the forecast value of the rival theory. I thus obtain

$$\begin{aligned} (\mathbf{I}_{TN} - \hat{\lambda}_0(\mathbf{I}_{\boldsymbol{T}} \otimes \mathbf{M}))\mathbf{Y} &= (\mathbf{I}_{TN} - \hat{\lambda}_0(\mathbf{I}_{\boldsymbol{T}} \otimes \mathbf{M}))(\mathbf{R}_0\boldsymbol{\gamma}_0 + \hat{\mathbf{Y}}_1^*\delta + \mathbf{e}_0) \\ \mathbf{Y}^*(\hat{\lambda}_0) &= \mathbf{R}_0^*(\hat{\lambda}_0)\boldsymbol{\gamma}_0 + \hat{\mathbf{Y}}_1^*(\hat{\lambda}_0)\delta + \mathbf{e}_0^*(\hat{\lambda}_0) \end{aligned} \tag{4.9}$$

which, using notation $\boldsymbol{\eta} = [\boldsymbol{\gamma}_0', \delta]'$, can be written as

$$\mathbf{Y}^*(\hat{\lambda}_0) = \mathbf{R}_{01}^{**}(\hat{\lambda}_0)\boldsymbol{\eta} + \mathbf{e}_{01}^{**}(\hat{\lambda}_0) \tag{4.10}$$

Finally, the test of the null model against the non-nested alternative is in terms of the null hypothesis *H0*: $\delta = 0$ against the alternative hypothesis *H1*: $\delta \neq 0$, i.e. a test of significance of whether the coefficient on the augmenting *H1* predictions in the augmented model is zero. I use instrument matrix $\mathbf{H}_{01}$, as defined previously, which gives the projection matrix $\mathbf{P}_{01}^{**} = [\mathbf{H}_{01}(\mathbf{H}_{01}'\hat{\boldsymbol{\Omega}}_{0\xi}\mathbf{H}_{01})^{-1}\mathbf{H}_{01}']$, so that the S2SLS / GMM estimator of $\boldsymbol{\eta}$ is

$$\hat{\boldsymbol{\eta}}_{S2SLS/GMM} = \left[\mathbf{R}_{01}^{**\prime}\mathbf{P}_{01}^{**}\mathbf{R}_{01}^{**}\right]^{-1}\mathbf{R}_{01}^{**\prime}\mathbf{P}_{01}^{**}\mathbf{Y}^*(\hat{\lambda}_0) \tag{4.11}$$

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

$$\hat{\mathbf{V}} = \left[(\mathbf{R}_{01}^{**\prime}\mathbf{H}_{01})(\mathbf{H}_{01}'\hat{\boldsymbol{\Omega}}_{0\xi}\mathbf{H}_{01})^{-1}(\mathbf{H}_{01}'\mathbf{R}_{01}^{**})\right]^{-1} = (\mathbf{R}_{01}^{**\prime}\mathbf{P}_{01}^{**}\mathbf{R}_{01}^{**})^{-1} \tag{4.12}$$

where $\hat{\mathbf{V}}(\hat{\delta})$ is the estimated variance of $\hat{\delta}$. The *J*-statistic is the square of the *t*-ratio associated with fitted values $\hat{\mathbf{Y}}_1^*(\hat{\lambda}_0)$ in the augmented model, i.e. under the null the following holds

$$\hat{J} = \frac{(\hat{\delta})^2}{\hat{\mathbf{V}}(\hat{\delta})} \to^d \chi_1^2 \tag{4.13}$$

with another way of defining *J*, equivalent to the above Wald statistics, being

$$\hat{J} = \frac{\hat{\delta}}{s.e.(\hat{\delta})} \to^d N(0,1) \tag{4.14}$$

In such expressions, $\chi_1^2$ and N(0,1) are the asymptotic reference distributions of the realised value of the *J*-statistic under each form of the test when the data are generated by the null.

The logic is that, if the coefficient on the added variable in the compound specification is significantly different from zero under the null model, then the null model is rejected; in other words, if the *J*-statistic is statistically significant (larger than the critical value for a conventional 5% probability in the relevant distribution), then this is evidence that the null model is rejected by the competing model. Moreover the test is asymmetric, meaning that rejecting *H0* does not imply that *H1* is true, so the procedure also involves testing the opposite case where *H1* is treated as maintained hypothesis; four outcomes are thus possible, i.e. rejection of *H0* by *H1*, rejection of *H1* by *H0*, rejection of both models, or neither model being rejected.

## 4.4 The empirical set-up

For the empirical application I use non-nested wage equations as in Table 3.1 (Model 4), Table 3.2 (Model 3) and Table 3.3 (Model 3) of section 3.6 of the previous chapter, respectively motivated by the wage curve, the Urban Economics (UE) theory and the New Economic Geography (NEG) theory.

A summary description of key variables (prior to natural log or spatial lag transformations) and data sources is given in section 3.4; a discussion of the literature, and details about econometric specifications, variables construction and spatial weighting matrices, are also given in the previous chapter.

The *J* -test technique is required to compare the wage curve and its UE or NEG rival as these are distinct models, with completely different provenance and contrasting hypotheses of wage determination. The wage curve is an empirical regularity, rather than a theoretically-derived relationship, between wage rates and unemployment rates. Meanwhile, UE and NEG are firmly grounded in economic theory, with microfoundations in the Cobb-Douglas production function for the competitive constant-returns final goods and services output, and the monopolistic competition and product variety theory of Dixit and Stiglitz (1977) for the imperfectly-substitutable increasing-returns intermediate market services inputs.

It should be noted that the hypotheses testing problem concerns (spatial) unemployment and its rival employment density or market potential propositions, while the supplementary labour efficiency variables are included to enhance the model (in the spirit of wage curve studies using micro data) or to operationalize the theory (in the case of UE or NEG).

## 4.5 Bootstrap inference in the *J*-test

### *4.5.1 Bootstrap design*

To implement a bootstrap test, I must generate *B* bootstrap samples, indexed by *z*, and recalculate $J$ for each of these simulated datasets, thus creating *B* bootstrap values of the test statistic, i.e. $\tilde{J}_z$. I adopt *B* = 999 which is a common and effective choice, offering proper levels of statistical power and clear test outcomes (MacKinnon, 2002).

There are two possible ways to produce bootstrap samples. One is the parametric bootstrap, where the vectors of errors are drawn from some distribution with specified moments; alternatively, bootstrap resampling (non-parametric bootstrap) has less demanding error distribution assumptions, since bootstrap errors are obtained by resampling from the null model's estimated residuals, and thus follow the empirical distribution function of these. Burridge and Fingleton (2010) actually adopt an intermediate approach whereby errors are simulated non-parametrically, using IV residuals from the first stage of the GMM estimation strategy as the building block for the residuals resampling, but a spatial autoregressive process is subsequently imposed in order to obtain SAR bootstrap errors. In this paper I specify a *parametric* bootstrap data generating process (DGP) assuming, for the simulated errors, a spatially autoregressive process with error components in space-time – the same structure as specified for the true DGP in eqs. (4.1a) and (4.1b). Below I describe the new parametric scheme for generating a null reference distribution for $J$ when non-nested hypotheses testing involves SARAR-RE models estimated by S2SLS / GMM.

The proposed approach requires simulating the finite-sample distribution of the test statistic using estimated parameters under the null hypothesis and SARAR-RE assumptions regarding spatial processes.

(a) First, I fit the null model to observed wage data to get parameter estimates $\hat{\iota}_0, \hat{\rho}_0, \hat{\alpha}_0, \hat{\boldsymbol{\beta}}_{\mathbf{0}}, \hat{\lambda}_0, \hat{\sigma}^2_{0u}, \hat{\sigma}^2_{0\nu}$. As in section 4.3, I use $\mathbf{Y}^*, \mathbf{Z}^*, \mathbf{X}^*, \mathbf{e}^*$ to denote variables after a spatial C-O transformation involving $\hat{\lambda}_0$. Hence, when the null is true, I assume that the following holds

$$\mathbf{Y}_0^* = \iota_0 \boldsymbol{\iota}_{TN} + \rho_0(\mathbf{I}_T \otimes \mathbf{W})\mathbf{Y}_0^* + \alpha_0 \mathbf{Z}_0^* + \mathbf{X}_0^* \boldsymbol{\beta}_0 + \mathbf{e}_0^* \tag{4.15}$$

(b) Secondly, I generate simulated wage data $\tilde{\mathbf{Y}}_0^*$ under the null via my estimates from Step (a) i.e. $\hat{\iota}_0, \hat{\rho}_0, \hat{\alpha}_0, \hat{\boldsymbol{\beta}}_{\mathbf{0}}, \hat{\lambda}_0, \hat{\sigma}_{0u}^2, \hat{\sigma}_{0\nu}^2$

$$\begin{aligned} \tilde{\mathbf{Y}}_0^* &= \hat{\iota}_0 \boldsymbol{\iota}_{TN} + \hat{\rho}_0(\mathbf{I}_T \otimes \mathbf{W})\tilde{\mathbf{Y}}_0^* + \alpha_0 \mathbf{Z}_0^* + \mathbf{X}_0^* \hat{\boldsymbol{\beta}}_0 + \tilde{\mathbf{e}}_0^* \\ \tilde{\mathbf{e}}_0 &= (\boldsymbol{\iota}_T \otimes (\mathbf{I}_N - \hat{\lambda}_0 \mathbf{M})^{-1})\tilde{\mathbf{u}}_0 + (\mathbf{I}_T \otimes (\mathbf{I}_N - \hat{\lambda}_0 \mathbf{M})^{-1})\tilde{\mathbf{v}}_0 \end{aligned} \tag{4.16}$$

where, at each $z$th replication, the error components $\tilde{\mathbf{u}}_0$ and $\tilde{\mathbf{v}}_0$ are randomly drawn from a Normal distribution with zero mean and estimated variances $\hat{\sigma}_{0u}^2$ and $\hat{\sigma}_{0\nu}^2$ as obtained from the null model

$$\tilde{\mathbf{u}}_0 \sim N(0, \hat{\sigma}_{0u}^2) \qquad \tilde{\mathbf{v}}_0 \sim N(0, \hat{\sigma}_{0\nu}^2) \tag{4.17}$$

It should be noted that the null in eqs. (4.15) and (4.16) is “true” in the sense that it refers to the model which is treated as the data-generating model, and tested against a non-nested alternative.

(c) The third stage of the bootstrap DGP involves fitting the alternative model (*H1*) to the simulated wage data $\tilde{\mathbf{Y}}_0^*$, in order to obtain predicted values from the *H1* model fitted to the null data

$$\hat{\tilde{\mathbf{Y}}}_1^* = E\left(\tilde{\mathbf{Y}}_0^*\right) \tag{4.18}$$

where $E$ denotes expectation.

(d) Next I apply a Cochrane-Orcutt transformation to the fitted values $\hat{\tilde{\mathbf{Y}}}_1^*$ from the *H1* model

$$\hat{\tilde{\mathbf{Y}}}_1^*(\hat{\lambda}_0) = (\mathbf{I}_{TN} - \hat{\lambda}_0(\mathbf{I}_T \otimes \mathbf{M}))\hat{\tilde{\mathbf{Y}}}_1^* \tag{4.19}$$

(e) Finally I fit the *H0* model augmented with $\hat{\tilde{\mathbf{Y}}}_1^*(\hat{\lambda}_0)$

$$\tilde{\mathbf{Y}}_0^* = \iota_0 \boldsymbol{\iota}_{TN} + \rho_0(\mathbf{I}_T \otimes \mathbf{W})\tilde{\mathbf{Y}}_0^* + \alpha_0 \mathbf{Z}_0^* + \mathbf{X}_0^* \boldsymbol{\beta}_0 + \hat{\tilde{\mathbf{Y}}}_1^*(\hat{\lambda}_0)\delta + \mathbf{e}_0^* \tag{4.20}$$

with the $z$th repetition of the $J$-statistics as given below

$$\tilde{J}_z = \frac{\hat{\delta}_z}{s.e.(\hat{\delta}_z)} \quad \text{or} \quad \tilde{J}_z = \frac{(\hat{\delta}_z)^2}{\text{var}(\hat{\delta}_z)} \tag{4.21}$$

(f) The procedure (Steps b-e) is repeated *B*=999 times to obtain a reference distribution (bootstrap distribution) for $J$ under the maintained model, hence subscript $z$ in eq. 4.21; this means generating 999 simulated datasets and recalculating the $J$ statistic for each of these, thus giving 999 bootstrap values for the test statistic. The empirical distribution function of the $\tilde{J}_z$ s, or bootstrap distribution, is an estimate of true cumulative distribution function of the $\tilde{J}_z$ s and can be denoted with $\hat{\tilde{F}}(J)$.

Results based on this bootstrap distribution will be assumed to hold, as these are based on the assumption that such a reference distribution is correct. In this case, our bootstrap method is an extension to panel data of the cross-sectional approach of Fingleton and Burridge (2010) and Burridge (2012), who also presented Monte Carlo evidence in support of the proposed procedure to construct a bootstrap distribution; this literature thus serves as a solid base for the bootstrap method in this paper.

As the $J$-test rejects in the upper tail of the null reference distribution, the bootstrap *P*-value is equal to the fraction of the time that the simulated $J$ values, $\tilde{J}_z$ s, are larger than the estimated $J$-statistic, $\hat{J}$; if the proportion of the $B$ bootstrap samples which exceed the observed test statistic is less than the chosen significance level, $\alpha$, then I reject the null hypothesis at that level. I thus have

$$\tilde{P}(\hat{J}) = \frac{1}{B}\sum_{z=1}^{B} I(\tilde{J}_z > \hat{J}) \tag{4.22}$$
$$\text{Reject null whenever } \tilde{P}(\hat{J}) < \alpha$$

where $I(\cdot)$ is an indicator variable which takes value one if the argument is true and zero otherwise. The bootstrap *P*-value makes sense intuitively; if very few of the $\tilde{J}_z$ s are greater than $\hat{J}$, then the probability of obtaining a test statistic which is more extreme than $\hat{J}$ will be low, thus $\tilde{P}(\hat{J})$ will be small, and the null model will be rejected. The bootstrap critical value is the $\tilde{J}_z$ value such that

$$1 - \hat{\hat{F}}(c_\alpha) = \alpha \tag{4.23}$$

where $\hat{\hat{F}}(J)$ is the bootstrap distribution; an equivalent way of expressing rejection of the null model is $\hat{J} > \tilde{c}_\alpha$, where $\tilde{c}_\alpha$ is number $(1-\alpha)(B+1)$ in the list of the $\tilde{J}_z$ s sorted from smallest to largest.

### 4.6 Results from the empirical application

Tables 3.1, 3.2 and 3.3 in the previous chapter summarise results from the wage curve, UE theory and NEG theory estimated individually; all models give acceptable outcomes and are suitable in understanding the geography of wages, as they account for almost the same proportion (around 70%) of local wage rate variance.

Tables 4.1 and 4.2 for wage curve *vs* NEG and Tables 4.3 and 4.4 for wage curve *vs* UE correspond to the compound specifications giving $\hat{J}$, the estimated $J$-statistics, in each case. Findings show that if the observed *J*-statistics is referred to the asymptotic $\chi_1^2$ - or N(0,1) - distribution then I would always reject the maintained hypothesis, concluding that either model is capable of falsifying its rival. This is because I have very small probabilities, e.g. for wage curve *vs* NEG I have Prob($\chi_1^2$ >11.3998) ~ 0 - or Prob(N(0,1)>3.3764) ~ 0 - for *H0*: Wage Curve and Prob($\chi_1^2$ >9.0850) ~ 0 - or Prob(N(0,1)>3.0141) ~ 0 - for *H0*: NEG.

**Table 4.1 Wage Curve *vs* NEG case. *H0*: Wage Curve, *H1*: NEG. Estimation of augmented Wage Curve with minimal IV set (*r*=0)**

| | Coefficient | se | *t*-stat |
|---|---|---|---|
| Constant | 0.1144 | 1.8028 | 0.0635 |
| Endog. Spatial Lag | -0.0889 | 0.0591 | -1.5067 |
| $(\mathbf{I}_T \otimes \mathbf{W})\text{Log}U$ | -0.1294 | 0.0218 | 5.9311*** |
| Fitted Values NEG | 0.9623 | 0.2850 | 3.3764*** |
| Log*T* | -0.0032 | 0.0126 | -0.2531 |
| Log*S* | 0.0068 | 0.0418 | 0.1637 |
| $\hat{J}$ -stat (chi2(1)) | | 11.3998*** | |
| $\lambda$ | | 0.7177 | |
| $\sigma_u^2$ | | 0.0047 | |
| $\sigma_v^2$ | | 0.0034 | |
| CORR [a] | | 0.6960 | |

[a] Correlation between observed and fitted values of **Log*Wage***.
* indicates significance at 10% level, ** indicates significance at 5% level, *** indicates significance at 1% level.

**Table 4.2 Wage Curve *vs* NEG case. *H0*: NEG, *H1*: Wage Curve. Estimation of augmented NEG model with minimal IV set (*r*=0)**

| | **Coefficient** | **se** | ***t*-stat** |
|---|---|---|---|
| Constant | -6.1400 | 3.8532 | -1.5935 |
| Endog. Spatial Lag | -0.2436 | 0.1097 | -2.2197** |
| $(\text{Log}MP\text{-}\rho(\mathbf{I}_T \otimes \mathbf{W})\text{Log}MP)$ | 0.1099 | 0.0338 | 3.2563*** |
| Fitted Values Wage Curve | 1.8478 | 0.6130 | 3.0141*** |
| Log$T$ | -0.0379 | 0.0255 | -1.4867 |
| Log$S$ | 0.0974 | 0.0686 | 1.4202 |
| $\hat{J}$ -stat (chi2(1)) | 9.0850*** | | |
| $\lambda$ | 0.7794 | | |
| $\sigma_u^2$ | 0.0050 | | |
| $\sigma_v^2$ | 0.0030 | | |
| CORR | 0.6707 | | |

**Table 4.3 The Wage Curve *vs* UE case. *H0*: Wage Curve, *H1*: UE. Estimation of augmented Wage Curve with minimal IV set (*r*=0)**

| | **Coefficient** | **se** | ***t*-stat** |
|---|---|---|---|
| Constant | -19.4853 | 7.3858 | -2.6382*** |
| Endog. Spatial Lag | -0.7277 | 0.2409 | -3.0207*** |
| $(\mathbf{I}_T \otimes \mathbf{W})\text{Log}U$ | -0.1666 | 0.0265 | -6.2944*** |
| Fitted Values UE | 4.0495 | 1.1628 | 3.4826*** |
| Log$T$ | -0.1444 | 0.0533 | -2.7092*** |
| Log$S$ | 0.4277 | 0.1610 | 2.6561*** |
| $\hat{J}$ -stat (chi2(1)) | 12.1283*** | | |
| $\lambda$ | 0.5877 | | |
| $\sigma_u^2$ | 0.0102 | | |
| $\sigma_v^2$ | 0.0047 | | |
| CORR | 0.5235 | | |

**Table 4.4 Wage Curve *vs* UE case. *H0*: UE, *H1*: Wage Curve. Estimation of augmented UE model with minimal IV set (*r* = 0)**

| | **Coefficient** | **se** | ***t*-stat** |
|---|---|---|---|
| Constant | -5.3396 | 3.4746 | -1.5368* |
| Endog. Spatial Lag | -0.2678 | 0.1116 | -2.4001*** |
| $(\text{Log}E\text{-}\rho(\mathbf{I}_T \otimes \mathbf{W})\text{Log}E)$ | 0.0190 | 0.0041 | 4.6895*** |
| Fitted Values Wage Curve | 1.8291 | 0.5538 | 3.3031*** |
| Log$T$ | -0.0463 | 0.0234 | -1.9798** |
| Log$S$ | 0.0541 | 0.0662 | 0.8165 |
| $\hat{J}$ -stat (chi2(1)) | 10.9105*** | | |
| $\lambda$ | 0.6962 | | |
| $\sigma_u^2$ | 0.0058 | | |
| $\sigma_v^2$ | 0.0033 | | |
| CORR | 0.6437 | | |

Parametric bootstrap results for wage curve *vs* NEG are reported in Tables 4.5 and 4.6. The top row of Table 4.5 refers to the assumptions of $r$=0 (minimal instrument set) and Normally distributed error components. Looking at this baseline scenario, the bootstrap shows that *H0*: NEG is discredited by the wage curve, since the fraction of times that the bootstrap *J*-statistics exceed $\hat{J}$ is equal to Prob( $\tilde{J}_z$ >9.0850)=4/1000=0.004; this leads to reject the null and infer that, given the maintained market potential hypothesis, the spatially-lagged unemployment adds significantly to the explanatory capability of the null model. Meanwhile, for *H0*: Wage Curve, the simulated *P*-value is Prob( $\tilde{J}_z$ >11.3998) = 76/1000 = 0.076 or 7.6% which means that the maintained spatial unemployment hypothesis is not rejected at a conventional significance rate of 5%.

The rest of the results in Table 4.5 are given for different instrument choices, following Burridge and Fingleton (2010); in their cross-sectional investigation, they explore the effects on the performance of the bootstrap from changing $r \in (0,1,2)$ - i.e. from using, respectively, a minimal, intermediate and rich set - and find that size control is best achieved when the dimension of the IV set is as small as possible. Here the outcome is that, also for $r$=1 and $r$=2, the wage curve is capable of falsifying the NEG theory while the NEG theory is clearly rejected by the wage curve in the bootstrap *J*-test, with empirical sizes lining up with asymptotics as the IV set gets smaller.

In Table 4.6 I experiment with different error distribution assumptions, using Student's *t*(5) as in Burridge and Fingleton (2010); the bootstrap $J$-test selects the wage curve over market potential as the dominant model of wage determination regardless of whether error components are distributed as Normal or Student's *t(5).*

All in all, there are indications of poor results from the asymptotic use of the $J$ -test, as the wage curve is rejected asymptotically but not when the bootstrap is used. Moreover, reading across the robustness checks, bootstrap evidence is in support of the wage curve, showing rejection of *H0*: NEG but not rejecting when *H0*: Wage Curve. This is illustrated in Figure 4.1.

It is worth noting that, as shown by Burridge (2012), a feature of the spatial GMM estimator in the context of non-nested hypotheses testing is that in small samples it often delivers spatial parameter estimates that lie outside the invertibility region of the model; if this happens, the underlying rationale of the spatial model breaks down (LeSage and Pace, 2009, p 26) and constructing bootstrap samples becomes problematic. To overcome this practical obstacle and implement the spatial bootstrap, Burridge and Fingleton (2010) replace $\hat{\rho}$ and $\hat{\lambda}$ with 0.97 whenever the model's estimates exceed one; Burridge (2012) finds that the issue can be removed by the use of quasi-maximum likelihood (QML), although the problem of parameter space for $\rho$ and $\lambda$ remains for certain parameter combinations. In my empirical applications, the larger sample (spatial panel data) emerges as a clear advantage and the spatial error and lag correlation coefficients prove to be legitimate estimates, thus guaranteeing the invertibility of the model and the viability of the bootstrap.

Turning to parametric bootstrap results for wage curve *vs* UE as reported in Tables 4.7 and 4.8, the bootstrap method confirms the asymptotic result that the wage curve is capable of falsifying the UE theory and that the null that employment density adds no explanatory information given the wage curve is also rejected, as the bootstrap *P*-values associated with the fitted values from the UE model are very small (below 5%) is all cases.

**Table 4.5 Wage Curve *vs* NEG case, Bootstrap results based on Normal errors**

| | $\tilde{\mathbf{u}}_0 \sim N(0,\hat{\sigma}^2_{0u})$ | | | | $\tilde{\mathbf{v}}_0 \sim N(0,\hat{\sigma}^2_{0v})$ | | | |
|---|---|---|---|---|---|---|---|---|
| | *H0*: Wage Curve, *H1*: NEG | | | | *H0*: NEG, *H1*: Wage Curve | | | |
| $r$ | $\hat{J}$ | Mean $\tilde{J}_z$ | Var $\tilde{J}_z$ | $Prob\left(\tilde{J}_z > \hat{J}\right)$ | $\hat{J}$ | Mean $\tilde{J}_z$ | Var $\tilde{J}_z$ | $Prob\left(\tilde{J}_z > \hat{J}\right)$ |
| 0 | 11.399 | 4.73 | 19.52 | 76/1000=**0.076** | 9.085 | 1.13 | 2.23 | 4/1000=**0.004** |
| 1 | 11.829 | 4.94 | 21.36 | 84/1000=**0.084** | 13.245 | 0.72 | 0.94 | 0/1000=**0.000** |
| 2 | 10.319 | 4.72 | 20.49 | 111/1000=**0.111** | 7.136 | 0.84 | 1.41 | 2/1000=**0.002** |

Minimal ($r$=0), intermediate ($r$=1) and rich ($r$=2) IV set.

**Table 4.6 Wage Curve *vs* NEG case, Bootstrap results based on *t(5)* errors** [a]

| | $\tilde{\mathbf{u}}_0 \sim$ Student's $t(5)$ | | | | $\tilde{\mathbf{v}}_0 \sim$ Student's $t(5)$ | | | |
|---|---|---|---|---|---|---|---|---|
| | *H0*: Wage Curve, *H1*: NEG | | | | *H0*: NEG, *H1*: Wage Curve | | | |
| $r$ | $\hat{J}$ | Mean $\tilde{J}_z$ | Var $\tilde{J}_z$ | $Prob\left(\tilde{J}_z > \hat{J}\right)$ | $\hat{J}$ | Mean $\tilde{J}_z$ | Var $\tilde{J}_z$ | $Prob\left(\tilde{J}_z > \hat{J}\right)$ |
| 0 | 11.399 | 4.17 | 16.40 | 52/1000=**0.052** | 9.085 | 1.16 | 2.30 | 3/1000=**0.003** |
| 1 | 11.829 | 4.36 | 17.84 | 66/1000=**0.066** | 13.245 | 0.85 | 1.47 | 0/1000=**0.000** |
| 2 | 10.319 | 4.08 | 16.42 | 85/1000=**0.085** | 7.136 | 0.79 | 1.26 | 3/1000=**0.003** |

Minimal ($r$=0), intermediate ($r$=1) and rich ($r$=2) IV set.
[a] Following Burridge and Fingleton (2010).

**Table 4.7 Wage Curve *vs* UE case, Bootstrap results based on Normal errors**

| | $\tilde{\mathbf{u}}_0 \sim N(0,\hat{\sigma}^2_{0u})$ | | | | $\tilde{\mathbf{v}}_0 \sim N(0,\hat{\sigma}^2_{0v})$ | | | |
|---|---|---|---|---|---|---|---|---|
| | *H0*: Wage Curve, *H1*: UE | | | | *H0*: UE, *H1*: Wage Curve | | | |
| $r$ | $\hat{J}$ | Mean $\tilde{J}_z$ | Var $\tilde{J}_z$ | $Prob\left(\tilde{J}_z > \hat{J}\right)$ | $\hat{J}$ | Mean $\tilde{J}_z$ | Var $\tilde{J}_z$ | $Prob\left(\tilde{J}_z > \hat{J}\right)$ |
| 0 | 12.128 | 3.88 | 8.94 | 13/1000=**0.013** | 10.911 | 0.92 | 1.59 | 2/1000=**0.002** |
| 1 | 12.367 | 3.42 | 8.48 | 9/1000=**0.009** | 13.509 | 0.76 | 1.03 | 0/1000=**0.000** |
| 2 | 15.576 | 2.00 | 5.17 | 0/1000=**0.000** | 8.898 | 0.74 | 1.18 | 0/1000=**0.000** |

See Table 4.5.

**Table 4.8 Wage Curve *vs* UE case, Bootstrap results based on *t(5)* errors**

| | $\tilde{\mathbf{u}}_0 \sim$ Student's $t(5)$ | | | | $\tilde{\mathbf{v}}_0 \sim$ Student's $t(5)$ | | | |
|---|---|---|---|---|---|---|---|---|
| | *H0*: Wage Curve, *H1*: UE | | | | *H0*: UE, *H1*: Wage Curve | | | |
| $r$ | $\hat{J}$ | Mean $\tilde{J}_z$ | Var $\tilde{J}_z$ | $Prob\left(\tilde{J}_z > \hat{J}\right)$ | $\hat{J}$ | Mean $\tilde{J}_z$ | Var $\tilde{J}_z$ | $Prob\left(\tilde{J}_z > \hat{J}\right)$ |
| 0 | 12.128 | 3.47 | 10.42 | 18/1000=**0.018** | 10.911 | 1.09 | 2.00 | 0/1000=**0.000** |
| 1 | 12.367 | 2.80 | 8.23 | 6/1000=**0.006** | 13.509 | 0.91 | 1.58 | 0/1000=**0.000** |
| 2 | 15.576 | 1.73 | 4.83 | 1/1000=**0.001** | 8.898 | 0.93 | 1.62 | 1/1000=**0.001** |

See Table 4.6.

## 4.7 Monte Carlo experiments

### *4.7.1 MC design*

The experimental design for the Monte Carlo (MC) simulations is based on a format which is the professional standard for doing MC in spatial panel econometrics (e.g. Baltagi, Fingleton and Pirotte, 2014).

For each sample size the regressand $y_{ij}$ is generated from a model of the form

$$y_{it} = \iota + \rho \sum_{j=1}^{N} W_{ij} y_{jt} + \beta x_{it} + e_{it} \quad i = 1,\dots,N; \quad t = 1,\dots,T$$

where the error term $e_{it}$ follows a SAR process

$$e_{it} = \lambda \sum_{j=1}^{N} W_{ij} e_{jt} + \xi_{it}$$

and $\xi_{it}$ has an error component structure

$$\xi_{it} = u_i + v_{it}$$

with $u_i \sim iid(0, \sigma_u^2)$, $v_{it} \sim iid(0, \sigma_v^2)$. The assumed values for the error variances are taken as $(\sigma_u^2, \sigma_v^2) = (0.01, 0.01)$ in line with actual estimates in Tables 3.1-3.3.

There are two regressors, one for the null model ( $\mathbf{X}_0$ ) and one for the alternative model ( $\mathbf{X}_1$ ). Following Baltagi, Fingleton and Pirotte (2014), the (*i*,*t*)th value of $\mathbf{X}_0$ is generated according to

$$x_{it} = \delta x_{it-1} + \zeta_{it}$$

where $\delta$ =0.6, $\zeta_{it} \sim N(0,1)$ and $x_{i0} = 0$. The (*i*,*t*)th value of $\mathbf{X}_1$ is generated in the same way as that of $\mathbf{X}_0$ (Kelejian and Piras, 2015). In all experiments, once generated, the values of the regressors are held fixed in the MC trials.

For the regression coefficients in the null and alternative models I respectively assume $\beta_0 = -0.1$ and $\beta_1 = 0.1$ in line with typical wage curve and UE / NEG effects, however I also show results for $\beta_0 = -0.5$ ($\beta_1 = 0.5$) and $\beta_0 = -1.0$ ($\beta_1 = 1.0$) as in Kelejian and Piras (2015). The assumed values of the spatial autoregressive parameters are also taken as $\rho = (0.2, 0.6)$ and $\lambda = 0.40$ (positive spatial correlation) in line with Kelejian and Piras (2015).

Following Kelejian and Prucha (1999), the row-normalised weights matrix is defined as “5 ahead and 5 behind”, so that each element $W$ is directly related to the five ones immediately after and immediately before it, and is selected in a “circular” fashion, for example $W_N$ is directly related to $W_1$ and $W_{N-1}$ and similarly $W_1$ is directly related to $W_2$ and $W_N$. The weighting matrix in the regression model is assumed to be the same as that in the error process (Kelejian and Piras, 2015).

For all experiments, a total of 5000 simulations were performed, with $s$=100 (no. of MC $J$-test replications) and $m$=50 (no. of bootstrap replications within each MC trial). Kelejian and Piras (2015) use 2000 as the number of iterations needed to obtain a 95% confidence interval of length 0.019 on the size of a test, so that estimates can be viewed as being significantly different from the theoretical 5% level if outside the interval (0.041, 0.060).

The empirical size and power estimates for the $s$ MC samples are summarised in Table 4.9. Each estimate of the size of the test at some nominal significance level is the proportion of the $m$ bootstrap samples on which the null is true, i.e. $\mathbf{Y}$ is generated using $\mathbf{X}_0$ and $\beta_0$, but is (wrongly) rejected according to eq. (4.22). Similarly each estimate of the power of the test at some nominal significance level is the proportion of the $m$ bootstrap samples on which the alternative is true, i.e. $\mathbf{Y}$ is generated using $\mathbf{X}_1$ and $\beta_1$, and the null is (correctly) rejected according to eq. (4.22).

### *4.7.2 MC results*

True rejection frequencies are consistently above the 5% Type I error rate, supporting the case for the bootstrap to provide correct sizes and test outcomes. This is different from Kelejian and Piras (2015) who found empirical sizes for moderately-sized samples which are close to their asymptotic approximations. Similarly to their findings, Table 4.9 shows that lower size (nearer 5%), and higher power, are associated with larger assumed values of the spatial autoregressive coefficient.

Importantly, it appears that the size of the test slowly converges to the asymptotic size as sample size increases, but is still very different within the range of sample sizes in Table 4.9; thus, although the size of the test tends to improve in larger samples, evidence of the size inflation problem remains, justifying the use of bootstrapping.

With regard to power, as found by Kelejian and Piras (2015) for the case when the difference between null and alternative pertains to the matrix of regressors, the *J*-test presents very high power even for relatively small samples.

**Table 4.9 Empirical size and power estimates from Monte Carlo trials**

| | | Empirical Size | | Empirical Power | |
|---|---|---|---|---|---|
| | | **Mean**[27] | **Median** | **Mean** | **Median** |
| **N=49 T=4** | | | | | |
| $\rho$ = 0.2 | $\beta_0$ = -1.0 $\beta_1$ = 1.0 | 0.5258 | 0.54 | 0.8682 | 0.92 |
| | $\beta_0$ = -0.5 $\beta_1$ = 0.5 | 0.5150 | 0.52 | 0.9992 | 1.00 |
| | $\beta_0$ = -0.1 $\beta_1$ = 0.1 | 0.4874 | 0.48 | 0.9998 | 1.00 |
| $\rho$ = 0.6 | $\beta_0$ = -1.0 $\beta_1$ = 1.0 | 0.5156 | 0.52 | 1.0000 | 1.00 |
| | $\beta_0$ = -0.5 $\beta_1$ = 0.5 | 0.5064 | 0.52 | 1.0000 | 1.00 |
| | $\beta_0$ = -0.1 $\beta_1$ = 0.1 | 0.4852 | 0.45 | 1.0000 | 1.00 |
| **N=100 T=4** | | | | | |
| $\rho$ = 0.2 | $\beta_0$ = -1.0 $\beta_1$ = 1.0 | 0.5158 | 0.51 | 0.9658 | 1.00 |
| | $\beta_0$ = -0.5 $\beta_1$ = 0.5 | 0.4750 | 0.49 | 0.9874 | 1.00 |
| | $\beta_0$ = -0.1 $\beta_1$ = 0.1 | 0.4564 | 0.38 | 0.9998 | 1.00 |
| $\rho$ = 0.6 | $\beta_0$ = -1.0 $\beta_1$ = 1.0 | 0.4836 | 0.52 | 1.0000 | 1.00 |
| | $\beta_0$ = -0.5 $\beta_1$ = 0.5 | 0.4708 | 0.42 | 1.0000 | 1.00 |
| | $\beta_0$ = -0.1 $\beta_1$ = 0.1 | 0.4532 | 0.51 | 1.0000 | 1.00 |
| **N=200 T=4** | | | | | |
| $\rho$ = 0.2 | $\beta_0$ = -1.0 $\beta_1$ = 1.0 | 0.4778 | 0.49 | 0.9912 | 1.00 |
| | $\beta_0$ = -0.5 $\beta_1$ = 0.5 | 0.4682 | 0.48 | 0.9998 | 1.00 |
| | $\beta_0$ = -0.1 $\beta_1$ = 0.1 | 0.4520 | 0.42 | 1.0000 | 1.00 |
| $\rho$ = 0.6 | $\beta_0$ = -1.0 $\beta_1$ = 1.0 | 0.4722 | 0.48 | 1.0000 | 1.00 |
| | $\beta_0$ = -0.5 $\beta_1$ = 0.5 | 0.4662 | 0.46 | 1.0000 | 1.00 |
| | $\beta_0$ = -0.1 $\beta_1$ = 0.1 | 0.4336 | 0.44 | 1.0000 | 1.00 |

[a] Total simulations ($s*m$) = 5000.
$s$ = no. of MC *J*-test replications, $m$ = no. of BS replications within each MC trial.

[27] The *J*-test is repeated $s$=100 times (MC iterations) and, within each of these, $m$=50 bootstrap iterations are performed to obtain size and power for each of the MC iterations. This means I have a size distribution and a power distribution both based on 100 MC replications, so that the figures reported in Table 4.9 are means and medians of the size or power estimates in this distribution.

## 4.8 Conclusions

The chapter combines the important and long-standing problem of non-nested hypotheses testing with the need to model patterns of spatial interaction in a panel data context. My proposed procedure for panel models is adapted from the original formulation of the *J*-test by Davidson and MacKinnon (1981) and from the spatial fixed-effects extension of Kelejian and Piras (2015).

Moreover I recognise that the *J*-test can be substantially oversized in small samples, so I develop a new parametric bootstrap scheme for generating a valid null reference distribution when non-nested hypotheses testing involves SARAR-RE models estimated by S2SLS / GMM.

Findings reveal that, in finite samples, the null reference distribution of *J* can be far from $\chi_1^2$ or N(0,1), and that critical values based on asymptotic theory tend to incorrectly reject the null model too often; the bootstrap distribution, in contrast, provides a better approximation of the small-sample distribution for *J* under the null. I examine the improvement in the finite-sample performance of the test assuming Normal and non-Normal errors, and in all cases the bootstrap-adjusted test emerges as a reliable and effective procedure in controlling significance levels in small samples. Monte Carlo trials confirm, even for sample sizes as large as two-hundred, that asymptotics would be inadequate and that the bootstrap is needed for the *J*-test to work well.

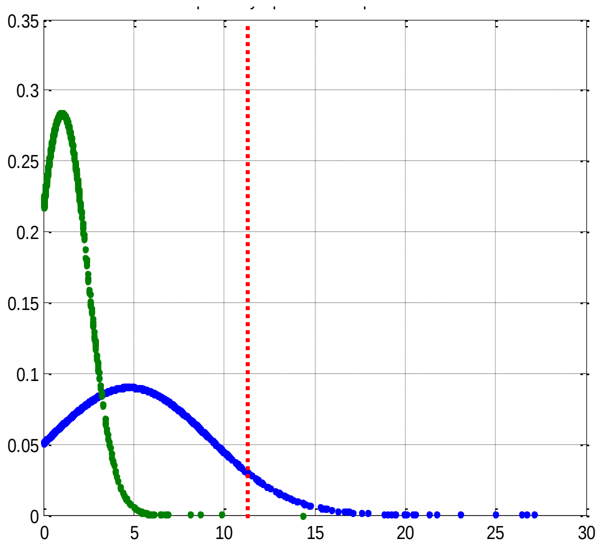
In terms of the empirical application, the *J*-test is implemented alternating between the wage curve and either NEG or UE as to which is treated as the maintained hypothesis. I find that the asymptotic *J*-test is too liberal and makes either model appear as capable of discrediting the maintained model, thus causing to always rejects the null. By constrast, the bootstrap helps correctly differentiate between the two hypotheses; it consistently points to the superiority of the wage curve over the NEG theory, while confirming the asymptotic result that the wage curve neither outperforms nor is outperformed by the UE theory. These *J*-test results are in line with asymptotic and bootstrap results from the Artificial Nesting Models in the previous chapter.

I can thus conclude that the unemployment rate kernel of the model motivated by the wage curve emerges as the prevailing explanation of local pay variation when tested against NEG but not against UE.

**Figure 4.1 Asymptotic Distribution *Z* (Green curve) & Bootstrap Distribution $\tilde{J}_z$ (Blue curve). Normal Errors, *r*=0 (999 samples)**

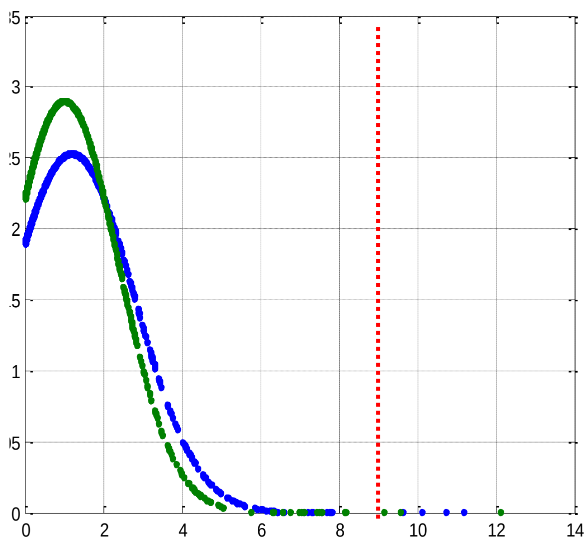
***H0:* Wage Curve, *HA:* NEG**

$\hat{J}$ =11.3998 ( $Z = \chi_1^2$ )

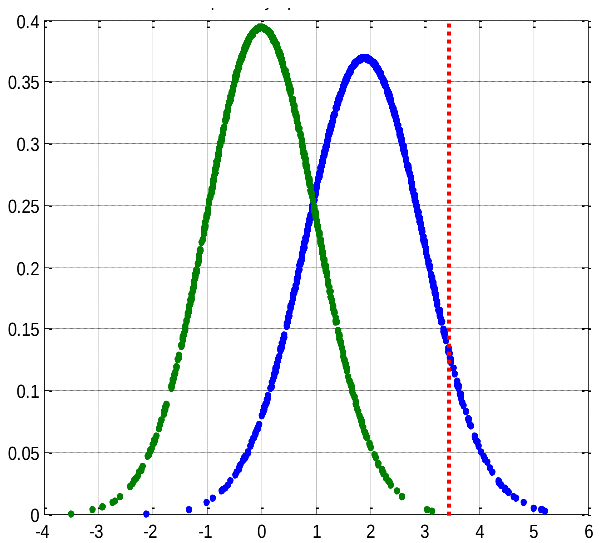



***H0:* NEG, *HA:* Wage Curve**

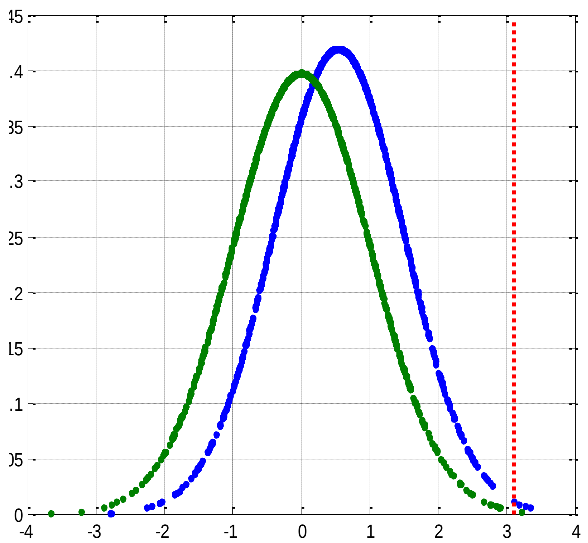
$\hat{J}$ =9.0850 ( $Z = \chi_1^2$ )




$\hat{J}$ =3.3764 ( $Z = N(0,1)$ )




$\hat{J}$ =3.0141 ( $Z = N(0,1)$ )



# Chapter 5

# Commuting Effects in Okun's Law among British Areas: Evidence from Spatial Panel Econometrics

## 5.1 Introduction

The linear, negative association linking real output growth to unemployment rate changes over the business cycle has been widely investigated since Arthur Okun's (1962, 1970) seminal work, which analyzed it using data for the U.S. in the years 1947-1960. The association in question, relating transitory movements in output and unemployment as measured by their year-on-year variations, is consistent with Okun's Difference Model, where first-differences represent the deviations of actual production and joblessness from their equilibrium trends. Another way of seeing the output-unemployment relationship, also studied by Okun (1962, 1970) and known as the Gap Model, is between the divergence of economic output from its potential or long-run level and the divergence of the joblessness rate from its non-inflationary or full-employment (NAIRU) level. The first-differences approach provides the base for the specification adopted in this paper.

Okun's finding of a 3:1 trade-off between economic growth and unemployment changes, often referred to as Okun's Law, has emerged as an empirical regularity predicting the magnitude of the reduction in unemployment from real GDP gains (or the costs in terms of higher unemployment of real GDP loss), and also how much demand stimulus is necessary to stabilise the joblessness rate. Specifically, the law envisages that for every 3% fall in output below its potential or long-term path the unemployment rate tends to rise by one percentage point (above its "natural", or NAIRU, level). This corresponds to a point estimate of around -0.3, i.e. a 1% fall in output from its trend yields an approximate 0.3 percentage point rise in the unemployment rate from its trend level. The less-than-proportionate increase in unemployment following an economic contraction, or equally the predicted slow employment response to an expansion in GDP, is due to labour-market stickiness. Firms tend to invest considerably in company-specific human capital and certain skills are in limited supply, meaning that they have more to lose by dismissing workers when faced with temporary downturns than by utilising labour less intensely in the short term. Equally, during upturns, firms may prefer to raise output via productivity gains instead of taking on new workers, which can result in sluggish labour market adjustments to positive demand shocks. It should be noted that the relationship can be, and indeed has been, explored in either causal direction depending on the empirical problem or policy question at hand; this paper analyses the responsiveness of unemployment to GDP performance, in line with models of unemployment typical of labour market research (for a survey, see Elhorst, 2003b).

While several tests of Okun's Law exist which are based on cross-country evidence (Knoester, 1986; Paldam, 1987; Moosa, 1997; Attfield and Silverstone, 1997, 1998; Lee, 2000; Freeman, 2001; Harris and Silverstone, 2001; Crespo-Cuaresma, 2003; Perman and Tavéra, 2005; Moazzami and Dadgostar, 2009), only recently have empirical studies started to estimate it using spatially-disaggregated data. The regional Okun's Law literature, mainly concerned with the responsiveness of output to unemployment, points to the existence of noticeable interregional differences in the size of the coefficient. Christopulous (2004), one of the main contributions, looks at thirteen Greek regions over the period 1971-1993; he reports slope values ranging from -0.37 to -1.70 in the areas where the empirical law holds, which tend to be areas with low levels of long-term unemployment; in contrast, areas where output and unemployment do not move together tend to be those where the majority of unemployment consists of people who have experienced a certain degree of deskilling. Adanu (2005) investigates Okun's Law for ten Canadian provinces during the period 1981-2001; his parameters vary from -0.30 to -2.14, with more negative values seen in areas with larger concentrations of skilled workers. Using 1980-2004 data for seventeen Spanish regions, Villaverde and Maza (2007, 2009) obtain regional estimates in the range of -0.32 to -1.55, and show that the law is stronger in areas where productivity growth is lower. The consensus is therefore that the output gains/costs of lower/higher unemployment are larger in some regions and smaller or even negligible in others. Two exceptions are the earlier studies by Freeman (2000) and Apergis and Rezitis (2003) who, focussing on eight regional economies respectively in the U.S. for the years 1958 to 1998 and in Greece over a similar time span 1960-1997, do not find clear evidence of spatial variability in the magnitude of Okun's coefficients.

While looking at the problem from a regional perspective, this body of work ignores the importance of controlling for spatial effects, and the implications for the strength and validity of Okun's relationship. To some extent, Kangasharju et al. (2012) represents an exception; the authors deal with the problem of cross-section dependence in their Okun's Law study of Finnish regions, however they do not treat it as the central theme of their paper but wash it out by taking output and unemployment series in deviation from their time means. By contrast, in labour market research, attempts have been made to rigorously apply a spatial economic/econometrics perspective to wage curve studies (Buettner, 1999; Baltagi et al., 2000; Longhi et al., 2006; Morrison et al., 2006; Elhorst et al., 2007), and have led to the conclusion that spatial effects matter and improve the explanatory power of the wage curve.

The present paper addresses this gap in the Okun's Law literature by exploring the question as to whether and to what extent there are spatial mechanisms involved in Okun's Law dynamics. For this purpose, I use data for the 128 NUTS3 regions of Great Britain over the period 1985-2011.

The paper is organised as follows: section 5.2 describes the model and data; section 5.3 is concerned with aspects of estimation; section 5.4 provides a discussion of results and, finally, section 5.5 summarises and concludes.

## 5.2 Model and data

### *5.2.1 Traditional Okun's Law specification*

Okun (1962, 1970) suggested two alternative forms of Okun's Law relationship, namely a gap model and a first-differences model. The former connects the deviation of actual output from its equilibrium or potential level to the unemployment rate gap; it therefore needs information about unemployment and output trends. Trend series are not directly observable, and there is no universal agreement on the optimal technique to estimate them, but any construction of these entails judgement. Moreover, the gap model should be preferred when the researcher is interested in inferences on time-series behaviour over the business cycle (Lee, 2000). By contrast the latter, relating real output growth to unemployment rate changes, has the advantage of not relying on approximations of the size of the gap. Thus, as is common practice in applied Okun's Law studies, and because my aim is not to document estimates under different approaches to trend estimation as in purely econometric exercises, I adopt the first-differences method of Okun's Law analysis.

The short-run relationship between output and unemployment as in Okun's (1962, 1970) first-differences version (see also Knoester, 1986) is given by the following expression

$$\Delta UN_{i,t} = \alpha + \beta \Delta GDP_{i,t} + e_{i,t} \quad , \quad \Delta GDP = \Delta \ln(\text{GDP}) \times 100 = \% \text{ growth rate} \qquad (5.1)$$

where $\Delta UN_{i,t}$ is the percentage point change in the local unemployment rate in region $i$ at time $t$ ($i$=1,…,$N$ with $N$=128 British NUTS3 areas, $t$=1985,…,2011 so that $T$=27), as constructed from claimant counts and working-age population data published by the United Kingdom's Office for National Statistics; $GDP_{i,t}$ is expressed as the natural logarithm of output (not in absolute terms), so that $\Delta GDP_{i,t}$ is the percent real growth rate of local economic activity, using Gross Value Added in basic constant (2006) prices as economic volume measure; and $\mathbf{e}_{i,t}$ is the error term, which in (non-spatial) Okun's Law studies is commonly modelled as satisfying the ordinary least squares assumptions of homoscedasticity and lack of autocorrelation. With regards to the structural parameters, $\alpha$ is the intercept, $\beta$ ( $\beta < 0$ ) is Okun's Law coefficient capturing the extent of the contemporaneous labour market reaction to short-term GDP fluctuations (estimated in the range -0.30 to -0.50), and the ratio $-\alpha / \beta$ indicates how fast economic activity has to grow in order to keep the unemployment rate stable. This basic regression equation can be augmented with other control variables which are commonly considered in the Okun's Law literature and justified on theoretical or empirical grounds.

In line with the dynamic version of Okun's Law (e.g. Chamberlin, 2011), I start by adding $\Delta UN_{i,t-1}$ to the right-hand side of eq. (5.1) in order to test whether current unemployment depends on its recent history. A significant influence of the unemployment rate in the preceding period would indicate the presence of rigidities and inertia in labour markets, causing delayed adjustments to workforce levels; it would also suggest the importance of path dependency and negative hysteresis (as discussed in Blanchard and Summers, 1987; Cross and Allan, 1988), thus reflecting persistent changes to the unemployment rate due to jobless workers permanently losing their skills or becoming inactive.

The extended specification, reflecting an unemployment-output relationship which is both contemporaneous and time-lagged, thus takes the form

$$\Delta UN_{i,t} = \alpha + \beta \Delta GDP_{i,t} + \gamma \Delta UN_{i,t-1} + e_{i,t} \tag{5.2}$$

As outlined in subsequent sections, the panel-data framework allows for unobserved or unmeasured time-invariant region-specific characteristics by means of fixed effects or of a composite error term structure incorporating random effects. These fixed-effects or random-effects vectors act as a catch-all for any causes of (time-constant) spatial heterogeneity. They include differences across regions in the sectoral composition of the local economic base, which can affect their relative ability to absorb demand shocks. For instance, output from the manufacturing and construction sectors is particularly sensitive to cyclical fluctuations. Moreover, employment in these industries consists in large part of temporary and contractual workers, who are easier to lay off when demand falters. Another way industrial structure can have supply-side effects on local unemployment is through the effectiveness of the skills-jobs matching process, which tends to be lower in regions specialized in agriculture and manufacturing (Taylor and Bradley, 1997; Elhorst, 2003b).

### *5.2.2 Labour market interactions*

While the regional dimension of Okun's Law is largely unexplored, the literature on regional unemployment disparities suggests that the geography of and interaction among regional economies are important drivers of labour-market outcomes. One of the first empirical papers to consider spatial variables in an unemployment determination model is Molho (1995), who looks at the geographical distribution of the joblessness rate across labour market areas within Great Britain in 1991. Starting from a standard regression of local unemployment on current and time-lagged local employment growth, he includes spatially lagged employment variables measuring demand changes in surrounding areas. He finds strong evidence of spillover effects from demand shocks both contemporaneously and after a lag, the former reflecting interregional trade links and the latter pointing to labour migration. The author also tests for the impact on each area's unemployment rate of that in neighbouring areas by incorporating the spatial lag of the dependent variable. His results, indicating a significant presence of spatial effects, are consistent with the transmission mechanism hypothesized by Burridge and Gordon (1981) and Taylor and Bradley (1983). These authors propose a balancing identity which relates regional unemployment changes to employment growth, labour force participation, migration and commuting, demonstrating the equilibrating effect of labour mobility on unemployment differentials. Such outcome arises because, as workers move from locations with spare capacity to locations with jobs surplus, local unemployment rates shift towards a new long-run steady state. A corollary to this is that labour-market developments are not confined to the local area but spill over to other areas as well, implying that regional unemployment will exhibit spatial autocorrelation. For instance, in a slack labour market, employers will find it less necessary to advertise vacancies outside their area and fewer workers from nearby regions will look to this area for jobs; thus, inward commuting flows will fall and labour-market conditions in contiguous areas will also tighten (see also Elhorst, 2003b). This explains why a region's unemployment rate tends to be higher/lower when surrounded by high-/low-unemployment regions.

Niebuhr (2003) follows this strand of analysis, looking at a sample of EU countries between 1986 and 2000. By means of spatial econometric techniques, she uncovers a strong degree of spatial linkages among European regional labour markets. In particular, the paper demonstrates that the evolution of a region's unemployment is considerably influenced by labour-market developments in surrounding regions, which lends support to the commuting hypothesis of Burridge and Gordon (1981) and other authors. She also tests for spatial dependence in the error term, and shows that factoring in spatial effects eliminates significant spatial residuals autocorrelation.

Building on this body of work, I introduce spatial effects in eq. (5.2), the dynamic counterpart of the standard Okun's Law relationship as given by eq. (5.1). Failure to account for labour market interactions can have serious consequences for the reliability of econometric results, as well as lead to an incorrect representation and understanding of the true causal forces at work. Specifically, neglecting spatial correlation in the variables of interest would cause biased and possibly inconsistent coefficient estimates, while leaving unobserved common factors (positive spatial residuals correlation) unmodelled would lead to reduced standard errors, inflated *t*-ratios and incorrect inference (Le Sage and Pace, 2009).

In the most complex of my specifications, spatial effects are in the form of spatial lags as well as spatially autoregressive error components, with regional heterogeneity modelled via random effects. The various spatial processes in this paper are summarized in the following regression equation; all of my estimated models are nested within this

$$\begin{aligned}\Delta UN_{i,t} &= \alpha + \rho \sum\nolimits_{j=1}^{N} W_{ij} \Delta UN_{j,t} + \beta \Delta GDP_{i,t} + \\ &\quad + \theta \sum\nolimits_{j=1}^{N} W_{ij} \Delta GDP_{j,t} + \gamma \Delta UN_{i,t-1} + e_{i,t} \\ e_{i,t} &= \lambda \sum\nolimits_{j=1}^{N} M_{ij} e_{j,t} + \xi_{i,t} \\ \xi_{i,t} &= \mu_i + v_{i,t}\end{aligned} \tag{5.3a}$$

where $e_{i,t}$ is the spatially dependent error term, and this is a function of $\xi_{i,t}$ which combines a time-invariant region-specific component $\mu_i \sim iid(0,\sigma_\mu^2)$ and a time-varying component $v_{i,t} \sim iid(0,\sigma_v^2)$ , respectively a random-effects vector and a disturbances or shocks vector (see section 2.3.4).

The fixed-effects counterpart to eq. (5.3a) can be formally expressed as

$$\begin{aligned} \Delta UN_{i,t} &= \alpha + \rho \sum\nolimits_{j=1}^{N} W_{ij} \Delta UN_{j,t} + \beta \Delta GDP_{i,t} + \\ &+ \theta \sum\nolimits_{j=1}^{N} W_{ij} \Delta GDP_{j,t} + \gamma \Delta UN_{i,t-1} + \mu_i + e_{i,t} \\ e_{i,t} &= \lambda \sum\nolimits_{j=1}^{N} M_{ij} e_{j,t} + v_{i,t} \end{aligned} \tag{5.3b}$$

### *5.2.3 Definition of Spatial Weight Matrices*

The *N* x *N* (standardized) spatial weights matrix $\mathbf{W}$, is used to construct spatial lags $(\mathbf{I}_T \otimes \mathbf{W})\Delta UN$ and $(\mathbf{I}_T \otimes \mathbf{W})\Delta GDP$, allows testing for the significance of spillover effects in labour-market and economic outcomes arising from workforce mobility. This takes the following form

$$\begin{aligned} W_{ij}^* &= \exp(-\hat{\tau}_i d_{ij}) \text{ for } i \neq j \\ W_{ij}^* &= 0 \qquad\qquad \text{for } i = j \\ W_{ij}^* &= 0 \qquad\qquad \text{for } \mathrm{d}_{ij} > 100\mathrm{km} \\ W_{ij} &= \frac{W_{ij}^*}{\sum_{j=1}^{N} W_{ij}^*} \end{aligned} \tag{5.4}$$

where $\hat{\tau}_i$ is specific to each area and calibrated on commuting flows (as explained in Appendix 4), with travel-to-work data taken from the UK's 2001 Census and converted from district to NUTS3 level, and $d_{ij}$ denotes the straight-line distance between any two areas *i* and *j*.

Regarding the distance threshold of 100km, I have taken this value from existing and well-established studies in the empirical regional economics literature; for example, Fingleton (2003) uses a similar specification of the **W** matrix (for Britain's local authority districts rather than NUTS3 areas) to explore the significance of increasing returns to labour productivity from employment density. Similarly, Lerbs and Oberst (2012) use a distance threshold of 90km in a four nearest neighbour inverse distance matrix. Sensitivity results are presented in Table 5.5. A discussion of the importance of using a cut-off distance is provided in section 2.3.3.

The $N$ x $N$ (standardized) spatial weights matrix **M** for the error process is based on a canonical contiguity specification, given by

$$
\begin{aligned}
&M_{ij}^{*} = 1 \text{ if } i \text{ and } j \text{ share a border} \\
&M_{ij}^{*} = 0 \text{ otherwise} \\
&M_{ij} = \frac{M_{ij}^{*}}{\sum_{j=1}^{N} M_{ij}^{*}}
\end{aligned}
\tag{5.5}
$$

With reference to the case at hand, the **W** matrix is constructed on commuting data because the main aim of this study is to explore the existence and significance of spatial effects due to workers mobility (labour market interactions), and a commuting-based spatial weights matrix enables me to explicitly test this proposition. For the error term, a spatial structure is typically imposed in order to capture common shocks as well as spatial autocorrelation in unobserved/ unmeasured causes of interregional heterogeneity (proximity effects, i.e. the fact that regions with similar socio-economic characteristics are typically close to each other); regarding **M**, a contiguity-based spatial weights matrix is a standard choice in spatial econometrics, and my preference for this functional form conforms to such custom.

## 5.3 Methodology

### *5.3.1 Instrumentation strategy*

It is important to observe that the model's variables may be jointly determined; $\mathbf{\Delta GDP}$ may be affected by two-way causation involving $\mathbf{\Delta UN}$, an aspect that has been neglected in the relevant Okun's Law literature to date, since unemployment is likely to cause (as well as be caused by) variations in demand. Similarly, with regard to $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{\Delta GDP}$, the construction of $\mathbf{W}$, which is based on commuting flows from the 2001 Census and thus postdates the dependent variable in some years, may introduce simultaneity bias and lead to inconsistent parameter estimates. I address these concerns using appropriate instruments to eliminate any correlation of Okun's variables with residuals, thereby ensuring that estimation results are accurate and reliable.

Thus, one of the elements of originality in this paper is the use of instrumental variables (IV) to correctly identify $\mathbf{\Delta GDP}$ and $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{\Delta GDP}$.

The instrumentation strategy is outlined in section 2.4 of the second chapter. More specifically, following the widely accepted and well established approach of Drukker, Egger and Prucha (2013), I specify an instruments set which includes the linearly independent columns of

$$\begin{gathered}\left[\mathbf{X}_f ;(\mathbf{I}_T \otimes \mathbf{W})\mathbf{X}_f ;(\mathbf{I}_T \otimes \mathbf{M})\mathbf{X}_f ;(\mathbf{I}_T \otimes \mathbf{W})(\mathbf{I}_T \otimes \mathbf{M})\mathbf{X}_f \right] \\ \mathbf{X}_f = \left[\Delta GDP_{i,t-1} ; \Delta UN_{i,t-1}\right]\end{gathered} \tag{5.6}$$

with spatial lags of first order as is common choice in applied spatial econometrics and in order to avoid over-parameterisation. This formulation relies on the variables in $\mathbf{X}_f$ being exogenous; both $\Delta UN_{i,t-1}$ and $\Delta GVA_{i,t-1}$ are pre-determined - i.e. predate the dependent variable, being lagged by one year - and can thus be treated as exogenous, although this property will be explicitly tested using appropriate diagnostics.

At a minimum, in my case the instruments should be a sub-set of eq. (5.6) containing the linearly independent columns of eq. (5.7) below (with spatial lags up to second order ($q$=2) as required to fulfil the rank condition for model over-identification)

$$\begin{gathered}\left[\mathbf{X}_f ; (\mathbf{I}_T \otimes \mathbf{W})\mathbf{X}_f ; (\mathbf{I}_T \otimes \mathbf{W})^2 \mathbf{X}_f \right] \\ \mathbf{X}_f = \left[\Delta GDP_{i,t-1} ; \Delta UN_{i,t-1}\right]\end{gathered} \tag{5.7}$$

I provide results both for this minimal representation of the IV set and for the extended expression/standard formulation with $q$=1.

### *5.3.2 Fixed effects or random effects*

Panel data estimation necessitates the selection between fixed effects and random effects. I make this choice on the basis of results from both specifications, taking into account theoretical coherence and empirical robustness, but my decision is also informed by the appropriate statistical devices which are available for this purpose; for instance, the Hausman statistics for random-effects consistency is normally used to this end, and the Sargan-Hansen instruments exogeneity test can help detecting misspecification and distinguishing between models. Therefore, I look for evidence in the data as to whether a random-effects model outperforms a fixed-effects model in this application.

### *5.3.3 Estimation*

All of my estimates from the models defined by eq. 5.3a and eq. 5.3b are derived from an instrumental-variables approach to satisfy orthogonality conditions and to achieve consistency, in contrast to existing spatial panel evidence on Okun's Law which is based on Maximum Likelihood and thus obtained under the assumption that all variables are exogenous.

Precisely, I carry out Two-Stage Least Squares (2SLS) estimation of Spatial Durbin Models (without spatially autoregressive errors) and Spatial Two-Stage Least Squares / Generalised Method of Moments (S2SLS / GMM) estimation.

The 2SLS technique is well known. With regard to the S2SLS / GMM procedure, I refer to the description of the procedure given in the introduction.

### *5.3.4 Measurement of direct, spillover and total impacts*

Le Sage and Pace (2009) point out that, when the spatial lags of the regressand and regressor are present in a model, the true total effect on a dependent variable (here $\mathbf{\Delta UN}$) of a unit change in an explanatory variable (here $\mathbf{\Delta GDP}$) – that is, the true partial derivative of the expected value of $\mathbf{\Delta UN}$ with respect to $\mathbf{\Delta GDP}$ (i.e. $\partial E(\mathbf{\Delta UN})/\partial \mathbf{\Delta GDP}$) - is not the same as the estimated regression coefficient $\hat{\beta}$; it also captures spatial linkages and simultaneous feedbacks passing through the dependence system, thus leading to a total effect which typically differs from $\hat{\beta}$ and which can be separated into a direct (own-region) effect and an indirect (spatial, spillover) effect.

Eq. (5.3a), which has both $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{\Delta UN}$ and $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{\Delta GDP}$ as determinants of $\mathbf{\Delta UN}$ (plus spatially autoregressive error components), accommodates regional interdependencies up and down the spatial network, thus expanding the information set for the *i*th region to include observations on the dependent and explanatory variables in other regions. The implication of including these spatial lags is that a unit change in $\mathbf{\Delta GDP}$ within a given area *i* at a given time *t* will directly affect $\mathbf{\Delta UN}$ in area *i* itself, but will also have an indirect effect on $\mathbf{\Delta UN}$ in all other areas which eventually impacts back to *i*. This is different from non-spatial linear regressions (based on the assumption of independence among cross-sectional units) where, in a given year *t*, $\partial E(\Delta UN_i)/\partial \Delta GDP_i = \beta$ for all *i* while $\partial E(\Delta UN_i)/\partial \Delta GDP_j = 0$ for $i \neq j$.

The proper interpretation of the marginal effects of $\mathbf{\Delta GDP}$ is derived from rearranging the following model, which is identical to eq. (5.3a) but expressed in terms of individual cross-sections, and taking expectations

$$\begin{aligned} E(\Delta UN_t) &= \alpha + \rho \mathbf{W}\Delta UN_t + \beta \Delta GDP_t + \theta \mathbf{W}\Delta GDP_t + \gamma \Delta UN_{t-1} = \\ &= (\mathbf{I}_N - \rho\mathbf{W})^{-1}\alpha + (\mathbf{I}_N - \rho\mathbf{W})^{-1}(\beta\Delta GDP_t + \theta\mathbf{W}\Delta GDP_t) \\ &\quad + (\mathbf{I}_N - \rho\mathbf{W})^{-1}\gamma\Delta UN_{t-1} \end{aligned} \tag{5.8}$$

where $|\rho| < 1$ , $\mathbf{I}_N$ is an $N$ x $N$ identity matrix and the Leontief Expansion $(\mathbf{I}_N - \rho\mathbf{W})^{-1}$ is equal to

$$(\mathbf{I}_N - \rho\mathbf{W})^{-1} = \mathbf{I}_N + \rho\mathbf{W} + \rho^2\mathbf{W}^2 + \rho^3\mathbf{W}^3 + \ldots \tag{5.9}$$

It follows that, at a given time *t*, the *N* x *N* matrix of partial derivatives of the expected value of $\mathbf{\Delta UN}$ in all regions with respect to $\mathbf{\Delta GDP}$ in all regions (i.e. $\partial E(\Delta UN_i)/\partial\Delta GDP_i$ for *i*=1,…,*n*) varies over *i* and can be illustrated to be equal to

$$\begin{bmatrix} \frac{\partial E(\Delta UN_1)}{\partial\Delta GDP_1} & \cdots & \frac{\partial E(\Delta UN_1)}{\partial\Delta GDP_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial E(\Delta UN_N)}{\partial\Delta GDP_1} & \cdots & \frac{\partial E(\Delta UN_N)}{\partial\Delta GDP_N} \end{bmatrix} = (\mathbf{I}_N - \rho\mathbf{W})^{-1} \begin{bmatrix} \beta & w_{12}\theta & \cdots & w_{1N}\theta \\ w_{21}\theta & \beta & \cdots & w_{2N}\theta \\ \vdots & \vdots & \ddots & \vdots \\ w_{N1}\theta & w_{N2}\theta & \cdots & \beta \end{bmatrix} \tag{5.10}$$

This matrix can be denoted by $\mathbf{S} = \partial E(\mathbf{\Delta UN})/\partial\mathbf{\Delta GDP} = (\mathbf{I}_N - \rho\mathbf{W})^{-1}\mathbf{C}$; the diagonal elements of $\mathbf{S}$ contain the direct impacts, and its off-diagonal elements represent indirect impacts (Le Sage and Pace, 2009).

More formally, the *average total effect* on $\mathbf{\Delta UN}$ of a unit change in $\mathbf{\Delta GDP}$ can be summarized by computing the row (or column) sum of partial derivatives contained in matrix $\mathbf{S}$ and then averaging over the *N* regions, as in

$$N^{-1}\sum_{ij}^{N}\frac{\partial E(\Delta UN_i)}{\partial\Delta GDP_j} = N^{-1}\boldsymbol{\iota}'\left[(\mathbf{I}_N - \rho\mathbf{W})^{-1}\mathbf{C}\right]\boldsymbol{\iota} \tag{5.11}$$

It is possible to distinguish the average total effects 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 the dependent variable $\mathbf{\Delta UN}$ in a single region *i* resulting from a unit change in the explanatory variable $\mathbf{\Delta GDP}$ across all *N* regions. The *average column effect* quantifies the *average total impact <u>from</u> 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 $\mathbf{\Delta UN}$ across all *N* regions resulting from a unit change in $\mathbf{\Delta GDP}$ in a single region *i*.

This average total effect can be partitioned into a direct component and an indirect component. The *average direct effect* is a scalar summary of the own-partial derivatives, each of these measuring the impact of a unit change in region *i*'s $\mathbf{\Delta GDP}$ on region *i*'s $\mathbf{\Delta UN}$. It is calculated as the average of the elements on the main diagonal of the $\mathbf{S}$ matrix, as in

$$N^{-1}\sum_{i}^{N}\frac{\partial E(\Delta UN_i)}{\partial \Delta GDP_i} = N^{-1}trace\left[(\mathbf{I}_N - \rho\mathbf{W})^{-1}\mathbf{I}_N\beta\right] \quad (5.12)$$

The *average indirect effect* is a scalar summary that corresponds to the cross-partial derivatives, each of these representing the response of region *i*'s $\mathbf{\Delta UN}$ to a unit change in $\mathbf{\Delta GDP}$ 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 $\mathbf{S}$. Results on the true Okun's Law coefficient, as obtained from the implementation of Le Sage and Pace's (2009) method, are presented in section 5.4.2.

## 5.4 Estimation results

### *5.4.1 Initial estimates*

This section presents results from my panel data analysis of Okun's Law for the 128 NUTS3 regions of Great Britain. Outcomes from numerous modelling solutions are provided in order to document biases due to model misspecification and to the omission of spatial patterns in space-time unemployment rate variations.

I start with fixed-effects (FE) and random-effects (RE) IV estimates from panel models with spatial lags of the dependent and independent variables but without interactions among errors, namely Spatial Durbin Models. **Table 5.1** illustrates initial estimates from baseline regressions, using an instruments set in its minimal specification (eq. 5.7).

**Table 5.1 Baseline regressions: Spatial Durbin Models with minimal IV set (eq. 5.7)**

| | 1 | 2 |
|---|---|---|
| *Estimation method* | **Fixed-Effects IV** | **Random-Effects IV** |
| Endogenous Spatial Lag ($\rho$) | **0.5770** | **0.4774** |
| (t-stat) | (5.29)*** | (4.14)*** |
| Real GDP growth rate (ΔGDP) ($\beta$) | **-0.0499** | **-0.0614** |
| (t-stat) | (-2.58)*** | (-2.46)*** |
| Spatial lag of ΔGDP ($\theta$) | **-0.0748** | **-0.0747** |
| (t-stat) | (-2.33)** | (-1.98)** |
| One-year lag of UN rate change ($\gamma$) | **0.1104** | **0.1480** |
| (t-stat) | (4.03)*** | (4.63)*** |
| Constant | | 0.2370 |
| (t-stat) | | (3.76)*** |
| *Diagnostics* | | |
| **Hausman test of regressors endogeneity** [a] | | |
| Chi-sq(2) statistic [P-value] | 14.07 [0.00] | 14.07 [0.00] |
| **First-stage F test of instruments relevance** [b] | | |
| F(6,3449) statistic [P-value] (ΔGDP) | 19.71 [0.00] | 19.71 [0.00] |
| F(6,3449) statistic [P-value] (Spatial lag of ΔGDP) | 41.47 [0.00] | 41.47 [0.00] |
| **Sargan-Hansen test of instruments orthogonality** | | |
| Chi-sq(2) statistic [P-value] | 4.75 [0.09] | 2.93 [0.23] |
| CORR [c] | 0.7671 | 0.7126 |
| No. regions | 128 | 128 |
| No. years (1985-2011) | 27 | 27 |

[a b] These diagnostics are common to both models.
[c] Correlation between observed and fitted values of **ΔUN**.
For the coefficient estimates, * indicates significance at 10% level, ** indicates significance at 5% level, and *** indicates significance at 1% level.

**Model 1** and **Model 2** are the closest approximations to a dynamic Okun's Law relationship with spatial effects, where joblessness rate changes are explained by time-lagged (local) unemployment rate changes, real GDP growth both locally and within commuting distance, and spatially-lagged (contemporaneous) unemployment rate changes.

The $\hat{\beta}$ coefficient of -0.05 (FE model) or -0.06 (RE model) is lower in absolute terms than Okun's value of -0.32, or than point estimates found elsewhere e.g. ranging from -0.17 to -0.24 in Crespo-Cuaresma (2003). Particularly, it is smaller than shown for the UK by studies which have investigated Okun's Law over various time periods using national data[27]. Nonetheless, with a *t*-ratio of -2.58 (FE model) or -2.46 (RE model), Okun's coefficient is highly significant, supporting the existence of a negative association between output and unemployment among the British regions. Previous studies of regional data, for instance Kangasharju et al. (2012), also found a smaller $\hat{\beta}$ coefficient than typically estimated from macro time-series data, and related this result to workers mobility and spatial linkages being more important across regions than across countries. The incorporation of spatial effects in the form of spatial lags of the dependent variable, $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{\Delta UN}$, and of the explanatory variable, $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{\Delta GDP}$, allows to explicitly test this hypothesis; given the definition of $\mathbf{W}$, the strength of spillovers between any two areas is inversely related to geographical distance and directly proportional to the intensity of the commuting links between them. I find a statistically significant and negatively signed between-area impact of real output changes on $\mathbf{\Delta UN}$, with the parameter on $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{\Delta GDP}$ equal to -0.07 in both models and with an associated *t*-ratio of -2.33 (FE model) or -1.98 (RE model). This validates the prediction that economic growth effects on $\mathbf{\Delta UN}$ are also due to real GDP performance in adjacent areas within commuting distance, rather than confined to the local area; therefore, Okun's Law papers should not focus only on the labour market responsiveness to output volume growth in a given location, but attention should be paid to how localised demand policies might influence workforce mobility and have employment effects outside administrative borders. Also for $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{\Delta UN}$, I find highly significant effects from nearby commuting areas, which corroborates the existence of interregional labour-market linkages.

Moreover I see strong relevance of the one-year lag of the dependent variable; in particular, the significant sign of the $\hat{\gamma}$ coefficient provides evidence of delayed adjustments to workforce levels and suggests the presence of negative

[27] For example, -0.34 in Knoester (1986), -0.36 in Paldam (1987), -0.38 in Moosa (1997), -0.69 in Attfield and Silverstone (1998), -0.77 in Freeman (2001) (see also Lee, 2000). For comparability with my result, these are the reciprocals of coefficients obtained from models with unemployment as right-hand-side variable.

hysteretic effects. Higher-order time lags are either insignificant or wrongly signed (compared to theoretical expectations as outlined in section 5.2.1), or both, and eliminating these does not modify my regression coefficient estimates, which implies that dynamic effects only accrue to the first period.

I next turn to the diagnostics. For the fixed-effects model (Model 1), the overidentifying restrictions test gives a Hansen-Sargan test statistic equal to 4.75, which when referred to the relevant $\chi^2$ distribution has an excedence probability of 0.09; this *P*-value is large enough to allow non-rejection of the null that instruments are exogenous, although it is relatively borderline. With regard to Model 2, the Hansen-Sargan test statistic is 2.93, with an associated *P*-value of 0.23, which indicates that orthogonality conditions hold strongly for the random-effects model. Also, in both cases, the joint *F* statistics from the first-stage regressions of 2SLS estimation are 19.71 for $\mathbf{\Delta GDP}$ and 41.47 for $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{\Delta GDP}$, all of which are extreme observations in the reference $F_{6,3449}$ distribution, which demonstrates that I do not have weak instruments.

Moreover, according to the Hausman test of regressors endogeneity, I can reject the null of exogeneity (i.e. the hypothesis that $\mathbf{\Delta GDP}$ and $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{\Delta GDP}$ are uncorrelated to residuals). Later estimations will actually produce test results which show absence of such correlation; however, it is worth noting that ultimately I select a specification with minimal IV set - which is the one that rejects the exogeneity hypothesis - so there is indeed a statistical need for using instruments.

Finally, I implement the Hausman consistency test which is a useful statistical tool to assess whether the difference in coefficients between the consistent fixed-effects model and the relatively more efficient random-effects model is significant; it uses the test statistic $\hat{H} = (\hat{\beta}_{RE} - \hat{\beta}_{FE})'(\hat{\Sigma}_{FE} - \hat{\Sigma}_{RE})^{-1}(\hat{\beta}_{RE} - \hat{\beta}_{FE})$, where the $\hat{\beta}s$ are the vectors of estimates from Model 1 and Model 2 and the $\hat{\Sigma}s$ are the respective covariance matrices. I find that $\hat{H}$ is not an extreme value with reference to the relevant $\chi^2$ distribution under the null, as the test statistic is equal to 10.35 which has a *P*-value of 0.0350; this implies that, at a conventional (though more liberal) 1%

level, I cannot reject the null that random effects are uncorrelated to the explanatory variables – in other words, statistical evidence is not enough to completely dismiss the estimates from Model 2.

**Table 5.2a Robustness checks: Spatial Durbin Models with extended IV set (eq. 5.6)**

| | 3a | 4a |
|---|---|---|
| *Estimation method* | **Fixed-Effects IV** | **Random-Effects IV** |
| Endogenous Spatial Lag ($\rho$) | **0.7741** | **0.6478** |
| (t-stat) | (14.19)*** | (9.69)*** |
| Real GDP growth rate (ΔGDP) ($\beta$) | **-0.0189** | **-0.0288** |
| (t-stat) | (-1.72)* | (-1.96)** |
| Spatial lag of ΔGDP ($\theta$) | **-0.0219** | **-0.0388** |
| (t-stat) | (-1.31) | (-1.80)* |
| One-year lag of UN rate change ($\gamma$) | **0.0731** | **0.1133** |
| (t-stat) | (4.38)*** | (5.48)*** |
| Constant | | 0.1108 |
| (t-stat) | | (3.78)*** |
| *Diagnostics* | | |
| **Hausman test of regressors endogeneity** | | |
| Chi-sq(2) statistic [P-value] | 2.29 [0.32] | 2.29 [0.32] |
| **First-stage F test of instruments relevance** | | |
| F(8,3447) statistic [P-value] (ΔGDP) | 26.02 [0.00] | 26.02 [0.00] |
| F(8,3447) statistic [P-value] (Spatial lag of ΔGDP) | 49.96 [0.00] | 49.96 [0.00] |
| **Sargan-Hansen test of instruments orthogonality** | | |
| Chi-sq(4) statistic [P-value] | 14.69 [0.01] | 14.82 [0.01] |
| CORR [a] | 0.8611 | 0.8318 |
| No. regions | 128 | 128 |
| No. years (1985-2011) | 27 | 27 |

[a] Correlation between observed and fitted values of **ΔUN**.
For the coefficient estimates, * indicates significance at 10% level, ** indicates significance at 5% level, and *** indicates significance at 1% level.

At this point in my analysis it seems that the fixed-effects model fits the data well and offers convincing results, but the random-effects model cannot be discarded because it produces equally satisfactory outcomes. Diagnostic evidence also suggests that the random-effects model should be retained: the Hausman FE vs. RE test indicates that random-effects estimates are consistent, as well as efficient; in addition, the Sargan-Hansen test shows that the instruments orthogonality conditions are strongly satisfied when random effects are used but less so under fixed effects, a possible consequence of some misspecification in the latter case. In order to shed further light on which model should be the preferred one, I next carry out robustness checks on my initial regressions.

**Model 3a** and **Model 4a** in **Table 5.2a** are obtained by applying an extended instruments set (eq. 5.6). Looking across Table 5.3a, it is apparent that the evidence points to the superiority of a random-effects model. Results from the fixed-effects model (Model 3a) become somewhat worse, with the $\hat{\theta}$ coefficient unexpectedly turning insignificant, while results from the random-effects model (Model 4a) remain statistically relevant. It should be noted that the Hausman regressors endogeneity test now confirms that the instrumented variables are not orthogonal to the error term, thus justifying the use of 2SLS estimation to guarantee consistency. Meanwhile, the first-stage $F$ statistic still leads me to conclude that instruments are relevant (i.e. jointly significant in identifying my endogenous regressors). With regard to the Hansen-Sargan overidentifying restrictions test, in both cases the test statistic is sufficiently small to infer that instruments are exogenous but, with the $P$-value just exceeding the 1% rate, hints at potential issues with these regressions. For the sensitivity checks in Table 5.2b, I thus use the minimal IV set, which proved to be valid according to diagnostics in Table 5.1, and I verify the robustness of estimates to the inclusion of dummy variables for the 1991-92 and 2008-09 recessions (similarly to Oberst and Oelgemöller, 2013).

**Table 5.2b Robustness checks: Spatial Durbin Models with crises dummies (eq. 5.7 for IVs)**

| | 3b | 4b |
|---|---|---|
| *Estimation method* | **Fixed-Effects IV** | **Random-Effects IV** |
| Endogenous Spatial Lag ($\rho$) | **0.8571** | **0.5345** |
| (t-stat) | (11.33)*** | (3.26)*** |
| Real GDP growth rate (ΔGDP) ($\beta$) | **-0.0266** | **-0.1068** |
| (t-stat) | (-1.11) | (-2.34)*** |
| Spatial lag of ΔGDP ($\theta$) | **-0.0135** | **-0.1207** |
| (t-stat) | (-0.35) | (-1.63)* |
| One-year lag of UN rate change ($\gamma$) | **0.0562** | **0.1149** |
| (t-stat) | (3.60)*** | (3.26)*** |
| 1991-92 Recession | **-0.0049** | **-0.2681** |
| (t-stat) | (-0.05) | (-1.60)* |
| 2008-09 Recession | **-0.1052** | **-0.6228** |
| (t-stat) | (-0.63) | (-2.15)*** |
| Constant | | 0.5361 |
| (t-stat) | | (2.21)** |
| *Diagnostics* | | |
| **Hausman test of regressors endogeneity** | | |
| Chi-sq(2) statistic [P-value] | 1.22 [0.54] | 1.22 [0.54] |
| **First-stage F test of instruments relevance** | | |
| F(6,3449) statistic [P-value] (ΔGDP) | 19.71 [0.00] | 19.71 [0.00] |
| F(6,3449) statistic [P-value] (Spatial lag of ΔGDP) | 41.47 [0.00] | 41.47 [0.00] |
| **Sargan-Hansen test of instruments orthogonality** | | |
| Chi-sq(2) statistic [P-value] | 17.62 [0.00] | 3.70 [0.16] |
| CORR [a] | 0.8647 | 0.6260 |
| No. regions | 128 | 128 |
| No. years (1985-2011) | 27 | 27 |

[a] Correlation between observed and fitted values of **ΔUN**.
For the coefficient estimates, * indicates significance at 10% level, ** indicates significance at 5% level, and *** indicates significance at 1% level.

From a comparison of **Model 3b** and **Model 4b** in **Table 5.2b** it emerges that results from a random-effects specification (Model 4b) are statistically relevant and economically meaningful while, in the fixed-effects specification (Model 3b), real GDP growth and its spatial lag are both statistically insignificant and well below theoretical expectations with regard to coefficient sizes. Importantly, the Sargan-Hansen test fails under fixed effects, suggesting instruments invalidity or other causes of model misspecification, whereas the overidentifying restrictions clearly hold in the context of IV estimation with random effects. The Hausman consistency test now gives a test statistic of 6.09, with a *P*-value of 0.73 which is well above a

10% significance level, strongly suggesting that random effects are uncorrelated to the explanatory variables.

**Table 5.3 Random-effects S2SLS / GMM estimation**

| | 5 | 6 |
|---|---|---|
| *Instrument set* | **Minimal (eq. 5.7)** | **Minimal (eq. 5.7)** |
| Endogenous Spatial Lag ($\rho$) | **0.3824** | **0.3319** |
| (t-stat) | (3.44)*** | (2.59)*** |
| Real GDP growth rate (ΔGDP) ($\beta$) | **-0.0609** | **-0.0652** |
| (t-stat) | (-2.66)*** | (-2.64)*** |
| Spatial lag of ΔGDP ($\theta$) | **-0.1119** | **-0.1232** |
| (t-stat) | (-3.00)*** | (-2.96)*** |
| One-year lag of UN rate change ($\gamma$) | **0.1615** | **0.1584** |
| (t-stat) | (6.36)*** | (6.23)*** |
| 1991-92 Recession | | **-0.1990** |
| (t-stat) | | (-1.54)* |
| 2008-09 Recession | | **-0.4203** |
| (t-stat) | | (-2.56)*** |
| Constant | 0.3009 | 0.3482 |
| (t-stat) | (3.33)*** | (3.14)* |
| *Spatial error process* | | |
| $\lambda$ | **0.5816*** ** | **0.6016*** ** |
| $\sigma_v^2$ | 0.1686 | 0.1728 |
| $\sigma_1^2 = \sigma_v^2 + T\sigma_u^2$ | 0.1519 | 0.1594 |
| CORR [a] | 0.8002 | 0.7782 |
| No. regions | 128 | 128 |
| No. years (1985-2011) | 27 | 27 |

[a] Correlation between observed and fitted values of **ΔUN**.
* is significance at 10% level, ** is significance at 5% level, *** is significance at 1% level.

All in all, estimation results from baseline and additional regressions support the adoption of a random-effects model, with a minimal form of the IV set fully satisfying the validity requirements (and noticeably outperforming its extended form on the basis of results from the Hansen-Sargan test). I thus select random effects because these worked well over various estimations, offering theoretically coherent and empirically solid outcomes, in contrast to fixed-effects evidence which proved less systematic and less acceptable than one would anticipate; available statistical tests such as the Hausman consistency test also support this decision. The final aspect in my spatial analysis is concerned with further developing my econometric model by allowing for spatially autoregressive (SAR) error components in a random-effects specification, estimated via S2SLS / GMM.

In **Table 5.3** (**Model 5** and **Model 6**), Okun's coefficient remains negative and significant, confirming the inverse response of unemployment rate changes to real output growth, and there is a statistically relevant estimate for the spatial autoregressive parameter $\lambda$ , which indicates significant 'nuisance' spatial dependence due to common shocks and omitted (positively) spatially autocorrelated explanatory variables.

Importantly, the inclusion of spatial dependence in the error term appears to improve the estimates. To begin with, results are more in line with expectations as the regression coefficient of own region's output growth is now lower, in absolute terms, than that of output growth in other regions within commuting distance; the theoretical argument for this is the presence of spatial dependence in labour markets from *systematic/substantive linkages*, here taking the form of commuting flows which cause labour-market outcomes in one place to depend partly on labour-market or business conditions elsewhere, and *non-systematic / nuisance linkages* which arise from inappropriate geographical delineation (spatial units used for data collection, i.e. formal / administrative areas, virtually never coincide with functional labour market areas, in other words output growth within commuting distance is a better approximation of the true extent of economic activity).

Also, compared to the counterpart models without SAR errors (2 & 4b), I find that $\mathbf{\Delta GDP}$ and particularly $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{\Delta GDP}$ become more strongly significant as *t*-ratios are higher; one reason is that the fit of the model is better when SAR errors are a component of the equation and thus standard errors are smaller. For $(\mathbf{I}_T \otimes \mathbf{W})\mathbf{\Delta UN}$, the absolute value of coefficient $\hat{\rho}$ is now lower, possibly because if $\mathbf{M}$-based interactions should be in the model and they are not then one will have omitted variables bias in the parameter estimates and this may be distorting $\hat{\rho}$ .

Models 5 and 6 give almost identical outcomes; however, since Model 5 has a slightly better fit, I take this as my final preferred specification. In the next section I use Model 5 as the point of departure to give a comprehensive and precise account of the validity and strength of Okun's Law, by capturing spatial effects in their full extent.

### *5.4.2 True Okun's Law coefficient: direct, spillover and total impacts*

Table 5.4 displays the true marginal effects relating to $\mathbf{\Delta GDP}$, i.e. the average total impact, and their average direct and indirect components. It should be noted that the direct and indirect effects associated with $\mathbf{\Delta GDP}$ are different from the values of, respectively, $\hat{\beta}$ and $\hat{\theta}$ obtained from Model 5 because they incorporate feedback loop effects, as implied by the Leontief Expansion (see eq. 5.10); these arise because any given area is considered a neighbour to its neighbour, so that shocks to the system propagate across neighbouring areas and eventually come back to the area they originated from.

The substantive result that emerges from this analysis is that the true Okun's Law coefficient amounts to -0.2798, which in absolute terms is very close to the 'empirical law', and this is primarily attributable to indirect effects. This means that, while the average direct impact coefficient of -0.0756 in Table 5.4 is statistically significant, what really counts for local labour market performance is real GDP growth in nearby areas within commuting distance, with a large part of the actual marginal effect of $\mathbf{\Delta GDP}$ on $\mathbf{\Delta UN}$ being due to spatial spillovers from workforce mobility – around 73%.

My conclusions regarding the validity and strength of Okun's Law for the British regions, based on the average total impact coefficient of -0.2798, are more in accord with existing evidence, as opposed to inference drawn from estimated regression coefficients $\hat{\beta}$ = -0.0609 and $\hat{\theta}$ = -0.1119 given by Model 5 (interpreted individually and without considering feedback loops). Spatial effects thus matter, and are responsible for the seemingly low impact of output on unemployment which is apparent from Model 5.

**Table 5.4 Direct, indirect and total effects of ΔGDP (Model 5)**

| W Matrix with distance threshold of 100 | Okun's Law Coefficient |
|---|---|
| **AVG DIRECT (OWN-REGION) EFFECTS** | **-0.0756** |
| (Bootstrapped *t*-ratio)[28] | (-3.19)*** |
| **AVG INDIRECT (SPATIAL, SPILLOVER) EFFECTS** | **-0.2042** |
| (Bootstrapped *t*-ratio) | (-5.18)*** |
| **AVG TOTAL EFFECTS** | **-0.2798** |
| (Bootstrapped *t*-ratio) | (-5.69)*** |
| **W Matrix with distance threshold of 150** | **Okun's Law Coefficient** |
| **AVG DIRECT (OWN-REGION) EFFECTS** | **-0.0754** |
| (Bootstrapped *t*-ratio) | (-3.55)*** |
| **AVG INDIRECT (SPATIAL, SPILLOVER) EFFECTS** | **-0.2039** |
| (Bootstrapped *t*-ratio) | (-5.16)*** |
| **AVG TOTAL EFFECTS** | **-0.2793** |
| (Bootstrapped *t*-ratio) | (-6.14)*** |
| **W Matrix with distance threshold of 50** | **Okun's Law Coefficient** |
| **AVG DIRECT (OWN-REGION) EFFECTS** | **-0.0763** |
| (Bootstrapped *t*-ratio) | (-3.47)*** |
| **AVG INDIRECT (SPATIAL, SPILLOVER) EFFECTS** | **-0.2056** |
| (Bootstrapped *t*-ratio) | (-4.71)*** |
| **AVG TOTAL EFFECTS** | **-0.2819** |
| (Bootstrapped *t*-ratio) | (-5.38)*** |

* is significance at 10% level, ** is significance at 5% level, *** is significance at 1% level.

For an explanation of why the regression coefficient of own region's output growth is lower in absolute terms than that of output growth in other regions within commuting distance, I refer the reader to the discussion of results from Table 5.3 in the previous sub-section. Moreover, my result is not exceptional, as authors who

---

[28] Bootstrap inference is based on comparison, in units of standard errors, of the observed average direct, indirect and total impact coefficients (those associated with and derived from the estimated regression parameters) with those obtained from the expected regression parameters under the null hypothesis that their true value is zero. The null-hypothesis coefficients are simulated by drawing at random from a Normal distribution with zero mean and a standard deviation (measure of their variability) equal to the standard error of the corresponding estimated coefficient; this is done in *Matlab* by computing $B./TRATS.*randn(size(B,1),1)+0$ for each of the performed one hundred iterations. If the observed direct, indirect and total impact coefficients are more than two standard errors away from what one would expect under the null, then the null would be rejected in favour of the conclusion that they differ significantly from zero.

have undertaken similar analyses of direct, spillover and total effects have reached analogous conclusions. Lerbs and Oberst (2012) investigate the influence of house prices on homeownership rates among German regions and see a total effect of -5.3 percentage points, of which only -1.7 points are due to own-region variations while -3.6 points correspond to the average cumulative indirect effect (68% of the overall impact). In a regional Okun's Law context, Oberst and Oelgemöller (2013) find that the magnitude of the total growth effect can be attributed to output variations in neighbouring areas for a proportion of almost 60%. In both cases, results are robust to alternative weighting matrices – namely a contiguity matrix and a four nearest neighbour inverse distance matrix with no distance threshold or with a threshold of 90km. While this evidence may be less striking than my estimate of 73%, it indicates a prominent role of interregional linkages, with spatial effects accounting for more than half of the total effect in both papers.

In order to show the robustness of results to the choice/specification of $\mathbf{W}$, I vary the distance threshold, which is a discretionary parameter though motivated by existing literature as explained in section 5.2.3. Results in Table 5.4 show that threshold values of 100km, 150km and 50km give broadly the same estimates of direct, indirect and total effects, which demonstrates that outcomes are robust to the use of different forms of the commuting-based $\mathbf{W}$ matrix; obviously, with a threshold of 50km, the spatial weight matrix is relatively more sparse (more of its elements are zero), so spillover and total effects become somewhat less significant.

## 5.5 Conclusions

This paper has been the first to analyze Okun's Law using a spatial panel approach on NUTS3 data for Great Britain. The relationship, expressing unemployment rate changes as an inverse function of real GDP growth as in Okun's seminal work, is adapted to include random or fixed effects, to control for regional heterogeneity, spatially autoregressive error components, to account for common shocks or omitted spatial autocorrelated factors, and spatial lags of the dependent and explanatory variables, to absorb spatial spillovers between neighbouring areas within commuting distance.

I find that the predicted negative relationship is corroborated by my data, and that the significance of Okun's Law coefficient is maintained under different model specifications and estimation methods, which demonstrates the robustness of my results.

The main conclusion I reach from estimating Okun's Law in a regional context is that the regression coefficient is lower than previously shown by cross-country evidence, and this is largely attributable to the spatial mechanisms which are at work in a small-area, short-distance scenario. Firstly, the inclusion of the spatial lags reveals that commuting and other spatial effects from geographical proximity / interregional linkages (as embodied in the **W** and **M** matrices, respectively) are relevant to a proper understanding of Okun's Law dynamics. Moreover, although Model 5 provided a satisfactory set of outcomes, my final preferred estimate (-0.28) is one which is obtained from taking full account of spatial effects, including the additional impact on any given area from changes cascading through the entire spatial hierarchy, and which points to a stronger Okun's relationship than inferred from the $\hat{\beta}$ and $\hat{\theta}$ parameters of Model 5. This provides further support for the important role of regional interdependencies in the context of Okun's Law.

# Chapter 6

# Conclusion

### 6.1 Summary and contributions

This final chapter presents, for the individual papers forming part of the thesis, an overall summary of the findings and contributions of these papers, their policy implications and limitations. I start by highlighting the elements of novelty in the work of each chapter, details of which are given below.

Chapter 2 focuses on estimation and prediction with a spatial panel data model; it analyses local wages for a representative sample of British towns during the historical period 1871-1906, using the Wage Curve as empirical base. It provides a contribution to the literature on regional disparities in economic resilience to recessions. Instead of assessing the local (direct plus indirect) impact of a one-time shock to wages in a given region, as in econometric impulse response functions (Fingleton, Garretsen and Martin, 2012), I adopt methods which allow for global shock transmission via a spatially autoregressive error process, so that the total impact also includes feedback effects, as well as for hysteretic changes to long-run growth paths. I find that a spatial panel data approach, implemented by means of a Generalized Method of Moments (GMM) procedure *à la* Kapoor, Kelejian and Prucha (2007) and Fingleton (2008a), is relevant for modelling wages and analysing resilience; the random-effects component of the error term captures any omitted (time-invariant) causes of interregional heterogeneity, also potentially subject to spatial autocorrelation, and by taking account of permanent cross-section variation it also picks up long-run effects. I show that coefficient estimates for the in-sample period are robust to alternative specifications of Spatial Weights Matrix **M**, as well as consistent since GMM estimation ensures that the outcomes are indeed consistent residuals; I also find evidence that spatial spillover effects (endogenous spatial lag and spatially correlated error components) should be incorporated in the earnings equation. Moreover, the model produces out-of-sample forecasts of local wages under a no-recession scenario, which represent the long-term underlying growth trend of wages against which their actual path can be compared, thus allowing a counterfactual analysis of economic resilience. In addition, I apply Goldberger's (1962) correction to both fitted and forecast values of local wages, in order to obtain in-sample and out-of-sample predictions that are unbiased and have the lowest prediction variance, which leads to better model's fit and improved counterfactuals.

Chapter 3 is concerned with model comparison and selection among rival non-nested specifications in a spatial panel context, and addresses the phenomenon of regional wage imbalances in Great Britain. This study is the first that aims to achieve some analytical tractability of non-nested models of wage determination for Great Britain using panel data, and the first to look at the (spatially-extended) Wage Curve *vis-à-vis* either New Economic Geography (NEG) or Urban Economics (UE). It is also the first wage curve application to Great Britain from a spatial econometrics perspective. In this chapter the Wage Curve is estimated by GMM as in the previous essay, but is directly confronted by each of two competing theories which also seek to explain spatial variation in wage rates. I evaluate the performance of the Wage Curve with respect to either UE or NEG via an Artificial Nesting Model including both unemployment and either employment density or market potential as regressors, to see whether the data generating process and thus the Wage Curve is encompassed by any of the alternative hypotheses. This is in line with the 'Inclusive Regression' approach suggested by Davidson & MacKinnon (1993) and Hendry (1995). The analysis allows drawing an initial conclusion as to whether the Wage Curve is indeed the superior proposition, as it retains its predictive ability in the presence of its rival, or proves to be falsified by either NEG or UE. The aim here is to establish whether the Wage Curve truly is an 'empirical law of economics', by testing the unemployment hypothesis under the direct challenge of the rival NEG and UE theories considered in turn, however the approach I propose is relevant for discriminating among competing models in all cases where one is not a reduced form of the other, and thus tests such as the Likelihood Ratio are not applicable. Results (when using bootstrap *P*-values) show that the NEG theory is dominated by the Wage Curve while the UE theory is a stronger contender since the employment-density hypothesis is not rejected by the unemployment-rate hypothesis; this is not surprising in a short-distance and small-scale spatial setting such as the one in this study, given the existing evidence on wage determination that market potential matters more at higher levels of spatial aggregation compared to employment density which is more relevant when analysing smaller areas.

In Chapter 4, I turn to a more formal and rigorous examination of the relative performance of the Wage Curve, and implement a random-effects version of the *J*-test developed by Kelejian and Piras (2015) for spatial panel data models with fixed effects. The technique is used in the context of the spatial panel estimation strategy adopted in the previous chapters; it is thus consistent with non-nested specifications which accommodate spatial dependence in both the dependent variable and the composite disturbance term (individual as well as remainder or disturbance error component), and appropriate for specifications with time-constant covariates and where $N$ is much larger than $T$ so that the bulk of variation is between rather than within. I also provide Monte Carlo evidence that there are distributional issues associated with an asymptotic use of the *J*-test procedure, principally because the asymptotic Standard Normal approximation of the reference distribution causes size distortion and overrejection of the null, therefore I apply the Bootstrap in order to improve the finite-sample properties of the spatial *J*-test and ensure reliable inference. This essay is an extension of the previous chapter but in itself provides a contribution to the spatial panel *J*-test literature. Results are in line with the 'Inclusive Regression' results in Chapter 3 (when using bootstrap *P*-values), indicating that in this context unemployment is a better predictor of wages than market potential while employment density is equally successful.

Chapter 5 addresses the need to bridge recent advances in spatial econometrics with economic problems usually examined with national data. It is the first essay in regional science to examine Okun's Law for Great Britain using a spatial panel approach. The basic specification used as a starting point for the spatial analysis is the relationship between real output growth and unemployment rate changes. This is extended to incorporate spatial lags of the dependent and explanatory variables, in order to test for the presence of commuting effects, and to accommodate spatial correlation involving error components. In this case, because I am less certain about the appropriatness of random effects and I compare Okun's Law coefficients under various estimation approaches, I fit spatial panel data models with either fixed effects or random effects to control for cross-region heterogeneity and find robust evidence of a significant Okun's Law coefficient. Moreover, as the presence of interaction terms $(\mathbf{I}_T \otimes \mathbf{W})\Delta UN$ and $(\mathbf{I}_T \otimes \mathbf{W})\Delta GDP$ complicates the interpretation of Okun's coefficient, I apply Le Sage and Pace's (2009) procedure for measuring the true effect and disentangling this into a direct/own-region impact and a indirect/spillover impact. Results show that spillover effects due to workforce mobility matter and have important implications for the validity and strength of the empirical law, with most of the actual marginal effect of $\Delta GDP$ on $\Delta UN$ being attributable to spillovers from labour commuting.

To summarise, as the increasing availability of spatial panel data calls for new modelling and analytical frameworks, the aim of this thesis is to develop econometric methods for estimation, prediction and inference with such data, and to apply these to specific problems in regional economics for small areas in Great Britain. The contribution comes from four individual but interrelated chapters. Chapter 2 addresses the question why some areas are more vulnerable to shocks than others, thus making a contribution to the economic resilience / economic greography literature; the aim is to examine the relative ability of local economies to preserve their long-run wage growth path in a recession, and it is the first paper to model the impact and transmission of shocks within a spatial panel data setting allowing for *global* spillovers as well as hysteretic effects. The paper is motivated by the fact that existing work is mainly conceptual and exploratory, so there is a call (e.g. Martin, 2012) for more formal and rigorous analyses. Chapter 3 evaluates the relative success of alternative, non-nested models of wage determination; one of these is the Wage Curve, which is usually referred to as an 'empirical law of economics', but there are other, distinct sources of regional wage variation, so the issue of which is the best explanation of spatial wage imbalances is important to the applied researcher; moreover, this is the first paper to look at economic agglomeration hypotheses for Great Britain from a spatial panel econometric perspective. Chapter 4 in an extension of the previous study. It devises a bootstrap *J*-test for spatial random-effects models to address the problem of non-nested hypotheses testing; it is the first paper to extend the spatial *J*-test literature to SARAR-RE models, and also to include parametric bootstrap inference. The motivation is that random effects have various benefits over fixed effects, and in many cases are the appropriate and preferred option, but the existing toolkit does not currently consider such models. Finally, Chapter 5 examines the role of spatial effects in the 'empirical law of economics' known as Okun's Law, since this relationship has mainly been considered at macro level with no attention to sub-national phenomena. It investigates the implications of regional interdependencies for the validity and strength of the unemployment-output relationship; moreover, it is one of few papers to revisit Okun's Law using spatial panel data, and the first to do so for Great Britain.

## 6.2 Limitations of this thesis and directions for future research

This section discusses the limitations of each essay and suggests a number of areas where further research could be usefully applied. With reference to Chapter 2, the study could be repeated using recent, rather than historical, data for the British areas. First, to the extent that small-area statistics are better available for the latest century than for older years, this could improve the quality of the variables, especially local unemployment which in this chapter is assumed to grow at the rate of national unemployment, and also the scope for investigating the actual drivers of the observed differences in economic resilience; secondly, it would be interesting to see whether the conclusions I have drawn for Great Britain from this historical analysis also hold in a contemporary scenario.

In Chapters 3 and 4 I have considered a single non-nested alternative and looked at specifications which differ in terms of regressors, so the bootstrap *J*-test procedure could be further extended to a problem involving two or more alternative hypotheses and/or different spatial weighting matrices. I thus suggest empirical applications involving wider model comparisons as a possible direction for future research. Moreover, although the novel random-effects estimation method by Piras (2013) is not thought to be particularly advantageous in this specific case for the reason noted in the chapter in question, I acknowledge that with more appropriate data it could indeed be useful, so I leave this as another possible future extension. It is also possible that the use of geographical units of different scale might lead to different cocnlusions with regard to the relative success of alternative regional agglomeration hypotheses, so another suggestion for researchers could be to carry out this analysis with larger spatial units.

With regard to Chapter 5, after this work was completed and submitted for publication, I became aware that a refinement of the spatial panel estimation strategy for dynamic models had been published by Baltagi, Fingleton and Pirotte (2014); it would be worthwhile to use this new method in future spatial panel studies of dynamic Okun's Law specifications, as it could help improve parameter estimates. Moreover, in this chapter I approximated the unemployment and output gaps by means of a first-differences transformation and estimated a model with unemployment as dependent variable, but other versions of Okun's Law are possible, namely a de-trending transformation and/or a production function looking at the impact on output of factors such as participation and activity rates, productivity, and population growth, as well as unemployment rates; therefore, alternative approaches can be adopted depending on the specific empirical problem or policy question at hand.

### 6.3 Policy implications

This section discusses how the models presented in each essay can be used for policymaking. The counterfactual analysis in Chapter 2 can tell why some areas display more or less resilience to recessions than others, with respect to one particular determinant which is economic structure (i.e. sectoral employment composition/concentration). This is important to understand how industrial policy can be most effective in protecting regions from shocks. Conditional on data availability, other policy levers that could be considered are export orientation, human capital, innovation rate, business and enterprise culture, institutional arrangements, and so on.

With regard to Chapters 3 and 4, the new results from spatial panel data on the Wage Curve for Great Britain can be used in policy modelling tools such as Computable General Equilibrium to inform the calibration of parameters. Moreover, although I chose to illustrate the *J*-test procedure in the context of wage equations, the approach is applicable in any circumstance in which the researcher is faced with a choice between non-nested specifications, or needs additional evidence to be confident in the robustness of the model. For instance, Holden and McGregor (1991) apply the *J*-test to the problem of finding the preferable deflator for nominal entities in the UK consumption function. In this particular case, the policy insight is that encouraging competitive labour markets / more elastic local labour supply (e.g. via greater mobility/ lower mobility costs), or non-traded intermediate producer service linkages, might prove most effective in boosting productivity and reducing disparities.

The point above for the Wage Curve can also be made for Okun's Law, which is a stylised figure used for macroeconomic simulations in a range of models, as well as an important forecasting and policymaking tool used to predict the benefits in terms of unemployment reduction of GDP growth. Moreover, findings from the spatial analysis in Chapter 5 suggest that the elasticity of unemployment to output, and thus the extent to which demand-stimulus interventions are effective, can be improved by policies aimed at boosting interregional labour mobility, for instance by investing in infrastructure or removing restrictions related to issues such as housing or job information.

## References


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

Adanu, K. (2005). A cross-province comparison of Okun's coefficient for Canada. *Applied Economics*, 37, 561–570.

Aldcroft, D. H. and Fearon, P. (1972). *British Economic Fluctuations 1790-1939*. Macmillan St. Martin's Press.

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

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. The Netherlands: Kluwer Academic Publishers.

Anselin, L. (2001). Spatial econometrics. In: B. H. Baltagi (Eds), *A Companion to Theoretical Econometrics* (pp. 310-330). Massachusetts: Blackwell Publishers.

Anselin, L. (2003). Spatial externalities, spatial multipliers, and spatial econometrics. *International Regional Science Review*, 26, 153-166.

Anselin, L. and Florax, R. (1995). *New Directions in Spatial Econometrics*. London: Springer.

Apergis, N. and Rezitis, A. (2003). An examination of Okun's law: evidence from regional areas in Greece. *Applied Economics*, 35, 1147–1151.

Arbia, G., Basile, R., and Piras, G. (2005). Using Spatial Panel Data in Modelling Regional Growth and Convergence. Istituto di Studi e Analisi Economica (ISAE) Working Paper 55.

Arrow, K. J. (1962). The economic implications of learning by doing. The Review of Economic Studies, 29, 155-173.

Arthur, W. B. (1989). Competing technologies, increasing returns and lock-in by historical events. *Economic Journal*. March 1990.

Arthur, W. B. (1990). Positive feedbacks in the economy. *Scientific American*. February 1990.

Arthur, W. B. (1994). *Increasing Returns and Path Dependence in the Economy*. Ann Arbor: University of Michigan Press.

Artis, M. J., Miguelez, E. and Moreno, R. (2012). Agglomeration economies and regional intangible assets: an empirical investigation. *Journal of Economic Geography*, 12, 1167-1189.

Attfield, C. and Silverstone, B. (1997). Okun's coefficient: a comment. *Review of Economics and Statistics*, 79, 326-329.

Attfield, C. and Silverstone, B. (1998). Okun's law cointegration and gap variables. *Journal of Macroeconomics*, 20, 626-637.

Baltagi, B. H. (2013). *Econometric Analysis of Panel Data.* 5th edition. New York: Wiley.

Baltagi, B. H. and Blien, U. (1998). The German wage curve: evidence from the IAB employment sample. *Economics Letters*, 61, 135–142.

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

Baltagi, B. H., Blien, U. and Wolf, K. (2009). New evidence on the dynamic wage curve for Western Germany: 1980–2004. *Labour Economics*, 16, 47–51.

Baltagi, B. H., Bresson, G. and Pirotte, A. (2012). Forecasting with spatial panel data. *Computational Statistics and Data Analysis*, 56, 3381-3397.

Baltagi, B. H., Egger, P. and Pfaffermayr, M. (2008). A Monte Carlo Study for pure and pre-test estimators of a panel data model with spatially autocorrelated disturbances. *Annales d'Économie et de Statistique*, 97/98, 11-38.

Baltagi, B. H., Fingleton, B. and Pirotte, A. (2014). Estimating and forecasting with a dynamic spatial panel data model. *Oxford Bulletin of Economics and Statistics*, 76, 112-138.

Baltagi, B. H. and Li, D. (2006). Prediction in the panel data model with spatial correlation: the case of liquor. *Spatial Economic Analysis*, 1, 75-185.

Baltagi, B. H., Song, S. H. and Koh, W. (2003). Testing panel data regression models with spatial error correlation. *Journal of Econometrics*, 117, 123-150.

Baltagi, B. & Liu, L. (2011). Instrumental variable estimation of a spatial autoregressive panel model with random effects. *Economics Letters*, 111, 135-137.

Barde, S. (2010). Increasing returns and the spatial structure of French wages. *Spatial Economic Analysis,* 5, 73–91.

Bell, B., Nickell, S. and Quintini, G. (2002). Wage equations, wage curves and all that. *Labour Economics,* 9, 341-360.

Blanchard, O. J. and Summers, L. H. (1987). Hysteresis in unemployment. *European Economic Review*, 31, 288-295.

Blanchflower, D. G. and Oswald, A. J. (1990). The wage curve. *Scandinavian Journal of Economics,* 92, 215-235.

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

Blanchflower, D. G. and Oswald, A. J. (1994b). Estimating a wage curve for Britain 1973-90. *The Economic Journal*, 104, 1025–1043.

Blanchflower, D. G. and Oswald, A. J. (1995). An introduction to the wage curve. *Journal of Economic Perspectives*, 9, 153–167.

Blanchflower, D. G. and Oswald, A. J. (2005). The wage curve reloaded. National Bureau of Economic Research Working Paper 11338, Cambridge MA.

Boschma, R. A. and Martin, R. L. (2009). *Handbook of Evolutionary Economic Geography*. Cheltenham: Edward Elgar.

Bowden, R. J. and Turkington, D. A. (1984). *Instrumental Variables*. Cambridge UK: Cambridge University Press.

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

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

Brakman, S., Garretsen, H. and 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, 777–800.

Brülhart, M. and Mathys, N. A. (2008). Sectoral agglomeration economies in a panel of European regions. *Regional Science and Urban Economics*, 38, 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, 47–67.

Burridge, P. (2012). Improving the *J*-test in the SARAR model by likelihood-based estimation. *Spatial Economic Analysis*, 7, 75-107.

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

Burridge, P. and Gordon, I (1981). Unemployment in the British metropolitan labour areas. *Oxford Economic Papers*, 33, 274-297.

Card, D. (1995). The wage curve: a review. *Journal of Economic Literature*, 33, 785-799.

Chamberlin, G. (2011). Okun's law revisited. *Economic and Labour Market Review*, 5, 104-132.

Christopulous, D. K. (2004). The relationship between output and unemployment: evidence from Greek regions. *Papers in Regional Science*, 83, 611–620.

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

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

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

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

Crespo-Cuaresma, J. (2003). Okun's law revisited. *Oxford Bulletin of Economics and Statistics*, 65, 439-451.

Cross, R. and Allan, A. (1988). On the history of hysteresis. In: R. Cross (Eds), *Unemployment, Hysteresis and the Natural Rate Hypothesis*. Oxford: Blackwell.

Cross, R. (1993). On the foundations of hysteresis in economic systems. *Economics and Philosophy*, 9, 53-74.

Cross, R., Grinfield, M. and Lamba, H. (2009). Hysteresis and economics. *Control Systems Magazine*, 29, 30–43.

Cross, R., Mcnamara, H. and Pokrovskii, A. (2010). Memory of recessions. Working Paper 10–09, Department of Economics, University of Strathclyde.

Dastoor, N. (1983). Some aspects of testing non-nested hypothesis. *Journal of Econometrics*, 21, 213-228.

David, P. A. (1985). Clio and the economics of QWERTY. *American Economic Review*, 75, 332–337.

David, P. A. (2005). Path dependence in economic processes: implications for policy analysis in dynamical systems contexts. In: K. Dopfer (Eds), *The Evolutionary Foundations of Economics* (pp. 151–194). Cambridge UK: Cambridge University Press.

David, P. A. (2007). Path dependence and historical social science: an introductory lecture. Stanford Institute for Economic Policy Research, Policy Paper 04-022.

Davidson, R. and MacKinnon, J. (1981). Several tests for model specification in the presence of alternative hypotheses. *Econometrica*, 49, 781–793.

Davidson, R. and MacKinnon, J. (1982). Some non-nested hypothesis tests and the relations among them. *Review of Economic Studies*, 49, 551–565.

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

Davidson, R. and MacKinnon, J. (2002). Bootstrap *J*-tests of non-nested linear regression models. *Journal of Econometrics*, 109, 167–193.

Delgado, M. and Stengos, T. (1994). Semiparametric specification testing of non-nested econometric models. *Review of Economic Studies*, 61, 291-303.

Dixit, A. and Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *American Economic Review*, 67, 297–308.

Drukker, D. M., Egger, P. and Prucha, I. R. (2013). On two-step estimation of a spatial autoregressive model with autoregressive disturbances and endogenous regressors. *Econometric Reviews*, 32, 686–733.

Duranton, G. and Puga, D. (2000). Diversity and specialisation in cities: why, where and when does it matter?. *Urban Studies*, 37, 533–555.

Duranton, G. and Puga, D. (2001). Nursery cities: urban diversity, process innovation, and the life cycle of products. *American Economic Review*, 91, 1454–1477.

Elhorst, J. P. (2003a). Specification and estimation of spatial panel data models. *International Regional Science Review*, 26, 244-268.

Elhorst, J. P. (2003b). The mystery of regional unemployment differentials; a survey of theoretical and empirical explanations. *Journal of Economic Surveys*, 17, 709-748.

Elhorst, J. P. (2010a). Spatial panel data models. In: M. M. Fischer & A. Getis (Eds) *Handbook of Applied Spatial Analysis* (pp. 377–407). Berlin: Springer.

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

Elhorst, J. P. (2012a). Matlab software for spatial panels. *International Regional Science Review*, 1-17.

Elhorst, J. P. (2012b). Dynamic spatial panels: models, methods, and inferences. *Journal of Geographical Systems*, 14, 5-28.

Elhorst, J. P. (2013). Spatial panel models. In: *Handbook of Regional Science*. Berlin: Springer.

Elhorst, J. P. (2014). *Spatial Econometrics: From Cross-sectional Data to Spatial Panels*. New York: Springer.

Elhorst, J. P., Blien, U. and Wolf, K. (2007). New evidence on the wage curve: a spatial panel approach. *International Regional Science Review*, 30, 173-191.

Fan, Y. and Li, Q. (1995). Bootstrapping *J* type tests for non-nested regression models. *Economic Letters*, 48, 107-112.

Feinstein, C. (1972). *National income, output and expenditure of the United Kingdom 1855-1965*. Cambridge UK: Cambridge University Press.

Fingleton, B. (2003). Increasing returns: evidence from local wage rates in Great Britain. *Oxford Economic Papers*, 55, 716–39.

Fingleton, B. (2005). Beyond neoclassical orthodoxy: a view based on the new economic geography and UK regional wage data. *Papers in Regional Science*, 84, 351–375.

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

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

Fingleton, B. (2008a). A generalized method of moments estimator for a spatial panel model with an endogenous spatial lag and spatial moving average errors. *Spatial Economic Analysis*, 3, 27-44.

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

Fingleton, B. (2009a). A generalized method of moments estimator for a spatial model with moving average errors, with application to real estate prices. In: G. Arbia & B. H. Baltagi (Eds), *Spatial Econometrics: Methods and Applications* (pp. 35-52). Heidelberg: Physica-Verlag.

Fingleton, B. (2009b). Prediction using panel data regression with spatial random effects. *International Regional Science Review*, 32, 195-220.

Fingleton, B. (2010). Predicting the Geography of House Prices. SERC Discussion Paper 45.

Fingleton, B. and Fischer, M. (2010). Neoclassical theory versus New Economic Geography. Competing explanations of cross-regional variation in economic development. *Annals of Regional Science*, 44, 467–491.

Fingleton, B., Garretsen, H. and Martin, R. L. (2012). Recessionary shocks and regional employment: evidence on the resilience of UK regions. *Journal of Regional Science*, 52, 109-133.

Fingleton, B. and Le Gallo, J. (2007). Finite sample properties of estimators of spatial models with autoregressive, or moving average, disturbances and system feedback. *Annales d'Économie et de Statistique*, 87/88, 39-62.

Fingleton, B. and Le Gallo, J. (2008). Estimating spatial models with endogenous variables, a spatial lag and spatially dependent disturbances: finite sample properties. *Papers in Regional Science*, 87, 319-339.

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

Fingleton, B. and Palombi, S. (2013) Spatial panel data estimation, counterfactual predictions, and local economic resilience among British towns in the Victorian era, *Regional Science and Urban Economics*, 43(4), 649-660.

Fingleton, B. and Palombi, S. (2013). The wage curve reconsidered: is it truly an empirical law of economics?. *Région et Développement*, 38, 49-92.

Freeman, D. G. (2000). Regional tests of Okun's law. *International Advances in Economic Research*, 6, 557–570.

Freeman, D. G. (2001). Panel tests of Okun's law for ten industrial countries. *Economic Inquiry*, 39, 511-523.

Friedman, M. (1993). The plucking model of business fluctuations revisited. *Economic Enquiry,* 31, 171-177.

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

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

Godfrey, L. G. (1983). Testing non-nested models after estimation by instrumental variables or least squares. *Econometrica*, 51, 355-366.

Godfrey, L. G. (1998). Tests of non-nested regression models: Some results on small sample behaviour and the bootstrap. *Journal of Econometrics*, 84, 59-74.

Godfrey, L. G. and Pesaran, M. (1983). Test of non-nested regression models. *Journal of Econometrics,* 21, 133-154.

Goldberger, A. S. (1962). Best linear unbiased prediction in the generalized linear regression model. *Journal of the American Statistical Association*, 57, 369-375.

Great Britain Historical GIS Project (2004). Great Britain Historical GIS. University of Portsmouth.

Greene, W. H. (2003). *Econometric Analysis*. New Jersey: Prentice Hall.

Hanson, G. (1998). Market potential, increasing returns, and geographic concentration. *Journal of International Economics*, 67, 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, 315–348.

Harris, R. and Silverstone, B. (2001). Testing for asymmetry in Okun's law: a cross-country comparison. *Economics Bulletin*, 5, 1–13.

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

Head, K. and Mayer, T. (2006). Regional wage and employment responses to market potential in the EU. *Regional Science and Urban Economics*, 36, 573-594.

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

Hill, E., Wial, H. and Wolman, H. (2008). Exploring regional economic resilience. Working Paper 2008-04, Institute of Urban and Regional Development, Berkeley.

Holden, D. and McGregor, P (1991). Determining the appropriate deflator of nominal magnitudes in the UK consumption function: a non-nested test approach. *Applied Economics*, 23, 781-790.

Holling, C. S. **(**1973). Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics*, 4, 1-23.

Holling, C. S. (1996). Engineering resilience versus ecological resilience. In: P. Schulze (Eds), *Engineering within Ecological Constraints* (pp. 31-44). Washington DC: National Academy Press.

Holling, C. S. (2001). Understanding the complexity of economic, ecological and social systems. *Ecosystems*, 4, 390-405.

Hunt, E. H. (1973). *Regional Wage Variations in Britain 1850-1914*. Oxford: Clarendon Press.

Huriot, J.-M. and Thisse, J.-F. (2000). *Economics of Cities*. Cambridge UK: 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, 414-420.

Kangasharju, A., Tavéra, C. and Nijkamp, P. (2012). Regional growth and unemployment: the validity of Okun's law for the Finnish regions. *Spatial Economic Analysis*, 7, 381-395.

Kapoor, M. H., Kelejian, H. H. and Prucha, I. R. (2007). Panel data models with spatially correlated error components. *Journal of Econometrics*, 140, 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, 3-11.

Kelejian, H. H. and Piras, G. (2011). An extension of Kelejian's *J*-test for non-nested spatial models. *Regional Science and Urban Economics*, 41,281-292.

Kelejian, H. H. and Piras, G. (2013). A *J*-test for panel models with fixed effects, spatial and time dependence. Regional Research Institute Working Paper 2013-03.

Kelejian, H. H. and Piras, G. (2015) An extension of the *J*-test to a spatial panel data framework, *Journal of Applied Econometrics*, DOI: 10.1002/jae.2425.

Kelejian, H. H. and 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, 99-121.

Kelejian, H. H. and Prucha, I. R. (1999). A generalised moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*, 40, 509-533.

Kelejian, H. H. and Prucha, I. R. (2004). Estimation of simultaneous systems of spatially interrelated cross sectional equations. *Journal of Econometrics*, 118, 27-50.

Kelejian, H. H. and Prucha, I. R. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*, 157, 53-67.

Kennedy, P. (2003). *A Guide to Econometrics*. Oxford: Blackwell.

Kennedy, S. and Borland, J. (2000). A wage curve for Australia?. *Oxford Economic Papers*, 52, 774–803.

Kim, C.-J. and Nelson, C. R. (1998). Friedman's plucking model of business fluctuations: tests and estimates of permanent and transitory components. *Journal of Money, Credit, and Banking*, 33, 317-334.

Knoester, A. (1986). Okun's law revisited. *Weltwirtschaftliches Archiv*, 122, 657-666.

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

Lee, C. H. (1979). *British Regional Employment Statistics 1841-1971*. Cambridge UK.

Lee, J. (2000). The robustness of Okun's law: evidence from OECD countries. *Journal of Macroeconomics*, 22, 331–356.

Lerbs, O. W. and Oberst, C. A. (2012). Explaining the spatial variation in homeownership rates: results for German regions. *Regional Studies*, 1, 1-22.

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

Longhi, S., Nijkamp, P. and Poot, J. (2006). Spatial heterogeneity and the wage curve revisited. *Journal of Regional Science*, 46, 707-731.

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

Marshall, A. (1920). Principles of Economics. London: Macmillan.

McAleer, M. (1995). The significance of testing empirical non-nested models. *Journal of Econometrics*, 67, 149–171.

MacKinnon, J., White, H. and Davidson, R. (1983). Test for model specification in the presence of alternative hypotheses: some further results. *Journal of Econometrics*, 21(1), 53-70.

MacKinnon, J. G. (2002). Bootstrap inference in econometrics. *Canadian Journal of Economics*, 35, 615–645.

Martin, R. L. (2012). Regional economic resilience, hysteresis and recessionary shocks. *Journal of Economic Geography*, 12, 1–32.

Martin, R. L. and Sunley, P. (2006). Path dependence and regional economic evolution. *Journal of Economic Geography*, 6, 395–437.

Martin, R. L. and Sunley, P. (2009). The place of path dependence in an evolutionary perspective on the economic landscape. In: R. Boschma & R. Martin (Eds), *Handbook of Evolutionary Economic Geography*. Cheltenham: Edward Elgar.

McGlade, J., Murray, R., Baldwin, J., Ridgway, K. and Winder, B. (2006). Industrial resilience and decline: a co-evolutionary approach. In: E. Garnsey & J. McGlade (Eds), *Complexity and Co-Evolution: Continuity and Change in Socio-Economic Systems* (pp. 147–176). Cheltenham: Edward Elgar.

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

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

Mitze, T. (2012). *Empirical Modelling in Regional Science*. Berlin: Springer.

Mizon, G. and Richard, J. (1986). The encompassing principle and its application to testing non-nested models”. *Econometrica*, 54, 657-678.

Moazzami, B. and Dadgostar, B. (2009). Okun’s law revisited: evidence from OECD countries. *International Business and Economics Research Journal*, 8, 21-24.

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

Moosa, I. A. (1997). A cross-country comparison of Okun’s coefficient. *Journal of Comparative Economics*, 34, 335–356.

Morrison, P. S., Papps, K. L. and Poot, J. (2006). Wages, employment, labour turnover and the accessibility of local labour markets. *Labour Economics*, 13, 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, 334–338.

Niebuhr, A. (2003). Spatial interaction and regional unemployment in Europe. *European Journal of Spatial Development*, 5, 1-26.

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

Nijkamp, P. and Poot, J. (2005). The last word on the wage curve?. *Journal of Economic Surveys*, 19, 421–450.

Oberst, C. A. and Oelgemöller, J. (2013). Economic growth and regional labour market development in German regions: Okun’s Law in a spatial context. FCN Working Paper 5/2013.

Okun, A. (1962). Potential GNP: its measurement and significance. *American Statistical Association, Proceedings of the Business and Economics Section*, 98–103.

Okun, A. (1970). *The Economics of Prosperity*. Brookings Institution, Washington DC.

Pace, R. K. and Zou, D. (2000). Closed-form maximum likelihood estimates of nearest neighbour spatial dependence. *Geographical Analysis*, 32, 154-172.

Paldam, M. (1987). How much does one percent of growth change the unemployment rate?. *European Economic Review*, 31, 306-313.

Partridge, M. (2005). Does income distribution affect U.S. state economic growth?. *Journal of Regional Science*, 45, 363–394.

Perman, R. and Tavéra, C. (2005). A cross-country analysis of the Okun's law coefficient convergence in Europe. *Applied Economics*, 37, 2501-2513.

Pesaran, M. H. (2004). General diagnostic tests for cross section dependence in panels. University of Cambridge, Department of Economics, Working Paper 0435.

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

Pimm, S. L. (1984). The complexity and stability of economic systems. *Nature*, 307, 321-326.

Piras, G. (2013). Efficient GMM estimation of a Cliff and Ord panel data model with random effects. *Spatial Economic Analysis*, 8, 370-388.

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

Roos, M. (2001). Wages and market potential in Germany. *Jahrbuch für Regionalwissenschaft* (*Review of Regional Research*), 21, 171-195.

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

Rostow, W. W. (1948). *British Economy of the Nineteenth Century*. Oxford: Oxford University Press.

Sato, Y. (2000). Search theory and the wage curve. *Economics Letters*, 66, 93–98.

Setterfield, M. (2010). Hysteresis. Working Paper 10-04, Department of Economics, Trinity College, Hartford, Connecticut.

Simmie, J. and Martin, R. L. (2010). The economic resilience of regions: an evolutionary approach. *Cambridge Journal of Regions, Economy and Society*, 3, 27–44.

Southall, H. R. (1986). Regional unemployment patterns among skilled engineers in Britain. *Journal of Historical Geography*, 12, 268-286.

Southall, H. R., Gilbert, D. R. and Gregory, I. (1999). *Great Britain Historical Database: Labour Markets Database, Statistics of Wages and Hours of Work, 1845-1913* [computer file]. Colchester, Essex: UK Data Archive [distributor], September 1999. SN: 3710.

Taylor, J. and Bradley, S. (1983). Spatial variations in the unemployment rate; a case study of North West England. *Regional Studies*, 17, 113-124.

Taylor, J. and Bradley, S. (1997). Unemployment in Europe: a comparative analysis of regional disparities in Germany, Italy and the UK. *Kyklos*, 50, 221-245.

Villaverde, J. and Maza, A. (2007). Okun's law in the Spanish regions. *Economics Bulletin*, 18, 1–11.

Villaverde, J. and Maza, A. (2009). The robustness of Okun's law in Spain, 1980-2004: regional evidence. *Journal of Policy Modelling*, 31, 289–297.

Walker, B., Gunderson, L., Kinzig, A., Folke, C., Carpenter, S. and Schultz, L. (2006). A handful of heuristics and some propositions for understanding resilience in socio-ecological systems. *Ecology and Society*, 11 (http://www.ecologyandsociety.org/vol11/iss1/art13/).

Appendix 1 (Chapter 2)

## Innovations and Error Variance-Covariance Matrices & Goldberger Correction

The vector of innovations for time *t* is

$$\boldsymbol{\xi}_t = \mathbf{u} + \mathbf{v}_t$$

while the vector of errors for time *t* is

$$\begin{aligned}
\mathbf{e}_t &= \lambda \mathbf{M}\mathbf{e}_t + \boldsymbol{\xi}_t = (\mathbf{I}_N - \lambda\mathbf{M})^{-1}\boldsymbol{\xi}_t \\
\mathbf{e}_t &= \mathbf{B}_N^{-1}\boldsymbol{\xi}_t = \mathbf{B}_N^{-1}(\mathbf{u} + \mathbf{v}_t)
\end{aligned}$$

where $\mathbf{B}_N = (\mathbf{I}_N - \lambda\mathbf{M})$. The full *NT* x 1vector of innovations is

$$\boldsymbol{\xi} = (\boldsymbol{\iota}_T \otimes \mathbf{u}) + \mathbf{v}$$

while the full *NT* x 1vector of errors is

$$\begin{aligned}
\mathbf{e} &= \lambda(\mathbf{I}_T \otimes \mathbf{M})\mathbf{e} + \boldsymbol{\xi} = (\mathbf{I}_{TN} - \lambda(\mathbf{I}_T \otimes \mathbf{M}))^{-1}\boldsymbol{\xi} \\
\mathbf{e} &= (\mathbf{I}_{TN} - \lambda(\mathbf{I}_T \otimes \mathbf{M}))^{-1}(\boldsymbol{\iota}_T \otimes \mathbf{u}) + (\mathbf{I}_{TN} - \lambda(\mathbf{I}_T \otimes \mathbf{M}))^{-1}\mathbf{v} = \\
&= (\boldsymbol{\iota}_T \otimes (\mathbf{I}_N - \lambda\mathbf{M})^{-1})\mathbf{u} + (\mathbf{I}_T \otimes (\mathbf{I}_N - \lambda\mathbf{M})^{-1})\mathbf{v} = \\
&= (\boldsymbol{\iota}_T \otimes \mathbf{B}_N^{-1})\mathbf{u} + (\mathbf{I}_T \otimes \mathbf{B}_N^{-1})\mathbf{v}
\end{aligned}$$

in which $\mathbf{I}_N$ is an *N* x *N* identity matrix, $\mathbf{I}_T$ is a *T* x *T* identity matrix, $\mathbf{I}_{TN}$ is a *NT* x 1 identity matrix, $\boldsymbol{\iota}_T$ is a *T* x 1 vector of ones, $\mathbf{u}$ is the *N* x 1 time-constant random-effects vector, and $\mathbf{v}$ is the *NT* x 1 time-varying shocks or disturbances vector.

Given eq. (2.8), the *NT* x *NT* innovations variance-covariance matrix $\boldsymbol{\Omega}_\xi$ is

$$\boldsymbol{\Omega}_\xi = E(\xi\xi') = \sigma_u^2(\mathbf{J}_T \otimes \mathbf{I}_N) + \sigma_v^2\mathbf{I}_{TN}$$

which is non-spherical and where $\sigma_1^2 = \sigma_v^2 + T\sigma_u^2$ and $\mathbf{J}_T = \boldsymbol{\iota}_T\boldsymbol{\iota}_T'$ is a *T* x *T* matrix of ones, while the *NT* x *NT* error variance-covariance matrix $\boldsymbol{\Omega}_e$ can be defined as

$$\mathbf{\Omega}_e = E(ee') = \mathbf{\Omega}_\xi\left[\mathbf{I}_{TN} - \lambda(\mathbf{I}_T \otimes \mathbf{M}))'(\mathbf{I}_{TN} - \lambda(\mathbf{I}_T \otimes \mathbf{M}))\right]^{-1}$$
$$\mathbf{\Omega}_e = E(ee') = \sigma_u^2\left[\mathbf{J}_T \otimes (\mathbf{B}_N'\mathbf{B}_N)^{-1}\right] + \sigma_v^2\left[\mathbf{I}_T \otimes (\mathbf{B}_N'\mathbf{B}_N)^{-1}\right]$$

- <u>Deriving the inverse of the NT x NT error variance-covariance matrix</u>, $\mathbf{\Omega}_e$.

Starting with

$$\mathbf{\Omega}_e = E(ee') = \sigma_u^2\left[\mathbf{J}_T \otimes (\mathbf{B}_N'\mathbf{B}_N)^{-1}\right] + \sigma_v^2\left[\mathbf{I}_T \otimes (\mathbf{B}_N'\mathbf{B}_N)^{-1}\right]$$

applying $(A\otimes B)+(C\otimes B)=(A+C)\otimes B$ and $k(A\otimes B)=(kA)\otimes B = A\otimes(k)B$

$$\mathbf{\Omega}_e = \left[\left(\sigma_u^2\mathbf{J}_T + \sigma_v^2\mathbf{I}_T\right)\otimes(\mathbf{B}_N'\mathbf{B}_N)^{-1}\right]$$

then taking the inverse and applying $((A)^{-1})^{-1} = A$, I have

$$\mathbf{\Omega}_e^{-1} = \left[\left(\sigma_u^2\mathbf{J}_T + \sigma_v^2\mathbf{I}_T\right)^{-1}\otimes(\mathbf{B}_N'\mathbf{B}_N)\right]$$

Collecting $\sigma_v^2$

$$\mathbf{\Omega}_e^{-1} = \left[\left(\sigma_v^2\left(\mathbf{I}_T + \frac{\sigma_u^2}{\sigma_v^2}\boldsymbol{\iota}_T\boldsymbol{\iota}_T'\right)\right)^{-1}\otimes(\mathbf{B}_N'\mathbf{B}_N)\right]$$

then applying $(\sigma^2 A)^{-1} = \dfrac{1}{\sigma^2}A^{-1}$, I have

$$\mathbf{\Omega}_e^{-1} = \left[\frac{1}{\sigma_v^2}\left(\mathbf{I}_T + \frac{\sigma_u^2}{\sigma_v^2}\boldsymbol{\iota}_T\boldsymbol{\iota}_T'\right)^{-1}\otimes(\mathbf{B}_N'\mathbf{B}_N)\right]$$

Applying $(\mathbf{I}_p + \mathbf{X}\mathbf{X}')^{-1} = \mathbf{I}_p - \mathbf{X}(\mathbf{I}_q + \mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ where $\mathbf{X}$ is $(p,q)$ and $\mathbf{X}'$ is $(q,p)$,

and with $\mathbf{X} = \dfrac{\sigma_u}{\sigma_v}\boldsymbol{\iota}_T$ and $\mathbf{X}' = \dfrac{\sigma_u}{\sigma_v}\boldsymbol{\iota}_T'$, I have

$$\mathbf{\Omega}_e^{-1} = \left[\frac{1}{\sigma_v^2}\left(\mathbf{I}_T - \frac{\sigma_u}{\sigma_v}\boldsymbol{\iota}_T\left(1 + \frac{\sigma_u^2}{\sigma_v^2}\boldsymbol{\iota}_T'\boldsymbol{\iota}_T\right)^{-1}\frac{\sigma_u}{\sigma_v}\boldsymbol{\iota}_T'\right)\otimes(\mathbf{B}_N'\mathbf{B}_N)\right]$$

$$\mathbf{\Omega}_e^{-1} = \left[\frac{1}{\sigma_v^2}\left(\mathbf{I}_T - \frac{\sigma_u^2}{\sigma_v^2}\boldsymbol{\iota}_T\left(1 + \frac{\sigma_u^2}{\sigma_v^2}\boldsymbol{\iota}_T'\boldsymbol{\iota}_T\right)^{-1}\boldsymbol{\iota}_T'\right)\otimes(\mathbf{B}_N'\mathbf{B}_N)\right]$$

$$\mathbf{\Omega}_e^{-1} = \left[\frac{1}{\sigma_v^2}\left(\mathbf{I}_T - \frac{\sigma_u^2}{\sigma_v^2}\boldsymbol{\iota}_T\left(1 + T\frac{\sigma_u^2}{\sigma_v^2}\right)^{-1}\boldsymbol{\iota}_T'\right) \otimes (\mathbf{B}_N'\mathbf{B}_N)\right]$$

where $p = T$ and $q = 1$ (as $\boldsymbol{\iota}_T$ is $(T,1)$, $\boldsymbol{\iota}_T'$ is $(1,T)$ and $\boldsymbol{\iota}_T'\boldsymbol{\iota}_T = T$ is $(1,1)$ so $\mathbf{I}_q = \mathbf{I}_1 = 1$).

Re-arranging, I have

$$\mathbf{\Omega}_e^{-1} = \left[\frac{1}{\sigma_v^2}\left(\mathbf{I}_T - \frac{\sigma_u^2}{\sigma_v^2}\boldsymbol{\iota}_T\boldsymbol{\iota}_T'\left(\frac{\sigma_v^2 + T\sigma_u^2}{\sigma_v^2}\right)^{-1}\right) \otimes (\mathbf{B}_N'\mathbf{B}_N)\right]$$

$$\mathbf{\Omega}_e^{-1} = \left[\frac{1}{\sigma_v^2}\left(\mathbf{I}_T - \frac{\sigma_u^2}{\sigma_v^2}\boldsymbol{\iota}_T\boldsymbol{\iota}_T'\frac{\sigma_v^2}{\sigma_v^2 + T\sigma_u^2}\right) \otimes (\mathbf{B}_N'\mathbf{B}_N)\right]$$

$$\mathbf{\Omega}_e^{-1} = \left[\frac{1}{\sigma_v^2}\left(\mathbf{I}_T - \frac{\sigma_u^2}{\sigma_1^2}\mathbf{J}_T\right) \otimes (\mathbf{B}_N'\mathbf{B}_N)\right]$$

which can also be written as

$$\mathbf{\Omega}_e^{-1} = \left[\frac{1}{\sigma_v^2}\left(\mathbf{I}_T - \frac{T\sigma_u^2}{\sigma_1^2}\frac{\mathbf{J}_T}{T}\right) \otimes (\mathbf{B}_N'\mathbf{B}_N)\right]$$

or equivalently

$$\mathbf{\Omega}_e^{-1} = \left[\frac{1}{\sigma_v^2}\left(\mathbf{I}_T - \frac{T\sigma_u^2}{\sigma_1^2}\bar{\mathbf{J}}_T\right) \otimes (\mathbf{B}_N'\mathbf{B}_N)\right]$$

○ <u>Deriving the NT x 1 vector of covariances between the prediction errors at location *i* and the sample errors</u>, $\boldsymbol{\omega}' = E(\mathbf{e}_{i,t+s}\mathbf{e}')$.

Starting with

$$E(\mathbf{e}_{t+s}\mathbf{e}') = E\left[\left(\mathbf{B}_N^{-1}(\mathbf{u} + \mathbf{v}_{t+s})\right)\left((\boldsymbol{\iota}_T \otimes \mathbf{B}_N^{-1})\mathbf{u} + (\mathbf{I}_T \otimes \mathbf{B}_N^{-1})\mathbf{v}\right)'\right]$$

and applying $I_T' = I_T$ and $(A \otimes B)' = (A' \otimes B')$, I have

$$E(\mathbf{e}_{t+s}\mathbf{e}') = E\left[\mathbf{B}_N^{-1}(\mathbf{u} + \mathbf{v}_{t+s})\left(\mathbf{u}'(\boldsymbol{\iota}_T' \otimes (\mathbf{B}_N^{-1})') + \mathbf{v}'(\mathbf{I}_T \otimes (\mathbf{B}_N^{-1})')\right)\right]$$

Re-arranging, I have

$$\begin{aligned} E(\mathbf{e}_{t+s}\mathbf{e}') = E[&\mathbf{B}_N^{-1}\mathbf{u}\mathbf{u}'(\boldsymbol{\iota}_T' \otimes (\mathbf{B}_N^{-1})') + \mathbf{B}_N^{-1}\mathbf{v}_{t+s}\mathbf{u}'(\boldsymbol{\iota}_T' \otimes (\mathbf{B}_N^{-1})') \\ &+ \mathbf{B}_N^{-1}\mathbf{u}\mathbf{v}'(\mathbf{I}_T \otimes (\mathbf{B}_N^{-1})') + \mathbf{B}_N^{-1}\mathbf{v}_{t+s}\mathbf{v}'(\mathbf{I}_T \otimes (\mathbf{B}_N^{-1})')] \end{aligned}$$

Taking expectations, I have

$$E(\mathbf{e}_{t+s}\mathbf{e}') = \mathbf{B}_N^{-1}E(\mathbf{uu}')(\boldsymbol{\iota}_T' \otimes (\mathbf{B}_N^{-1})') + \mathbf{B}_N^{-1}E(\mathbf{v}_{t+s}\mathbf{u}')(\boldsymbol{\iota}_T' \otimes (\mathbf{B}_N^{-1})') \\ + \mathbf{B}_N^{-1}E(\mathbf{uv}')(\mathbf{I}_T \otimes (\mathbf{B}_N^{-1})') + \mathbf{B}_N^{-1}E(\mathbf{v}_{t+s}\mathbf{v}')(\mathbf{I}_T \otimes (\mathbf{B}_N^{-1})')$$

where $E(\mathbf{uu}') = \sigma_u^2$, $E(\mathbf{v}_{t+s}\mathbf{u}') = 0$, $E(\mathbf{uv}') = 0$, $E(\mathbf{v}_{t+s}\mathbf{v}') = 0$.

Thus I have

$$E(\mathbf{e}_{t+s}\mathbf{e}') = \sigma_u^2 \mathbf{B}_N^{-1}(\boldsymbol{\iota}_T' \otimes (\mathbf{B}_N^{-1})')$$

and

$$\boldsymbol{\omega}' = E(\mathbf{e}_{i,t+s}\mathbf{e}') = \sigma_u^2 b_i (\boldsymbol{\iota}_T' \otimes (\mathbf{B}_N^{-1})')$$

where $b_i$ is the $i$th row of matrix $\mathbf{B}_N^{-1}$.

○ <u>Deriving Goldberger's (1962) BLUP correction</u> $\boldsymbol{\omega}'\boldsymbol{\Omega}_e^{-1}$.

Starting with

$$\boldsymbol{\omega}'\boldsymbol{\Omega}_e^{-1} = \frac{\sigma_u^2}{\sigma_v^2} b_i \left(\boldsymbol{\iota}_T' \otimes (\mathbf{B}_N^{-1})'\right)\left(\left(\mathbf{I}_T - \frac{T\sigma_u^2}{\sigma_1^2}\bar{\mathbf{J}}_T\right) \otimes (\mathbf{B}_N'\mathbf{B}_N)\right)$$

Applying $(A^{-1})' = (A')^{-1}$ and $(A \otimes B)(C \otimes D) = (AC \otimes BD)$, I have

$$\boldsymbol{\omega}'\boldsymbol{\Omega}_e^{-1} = \frac{\sigma_u^2}{\sigma_v^2} b_i \left(\boldsymbol{\iota}_T'\left(\mathbf{I}_T - \frac{T\sigma_u^2}{\sigma_1^2}\bar{\mathbf{J}}_T\right) \otimes (\mathbf{B}_N')^{-1}\mathbf{B}_N'\mathbf{B}_N\right)$$

Applying $A^{-1}A = I$ and $AI = A$, I have

$$\boldsymbol{\omega}'\boldsymbol{\Omega}_e^{-1} = \frac{\sigma_u^2}{\sigma_v^2} b_i \left(\left(\boldsymbol{\iota}_T' - \frac{T\sigma_u^2}{\sigma_1^2}\boldsymbol{\iota}_T'\right) \otimes \mathbf{B}_N\right)$$

where $\boldsymbol{\iota}_T'\mathbf{I}_T = \boldsymbol{\iota}_T'$ and $\boldsymbol{\iota}_T'\bar{\mathbf{J}}_T = \boldsymbol{\iota}_T'\left(\frac{1}{T}\boldsymbol{\iota}_T\boldsymbol{\iota}_T'\right) = \frac{1}{T}\boldsymbol{\iota}_T'\boldsymbol{\iota}_T\boldsymbol{\iota}_T' = \frac{1}{T}(\boldsymbol{\iota}_T'\boldsymbol{\iota}_T)\boldsymbol{\iota}_T' = \frac{1}{T}(T)\boldsymbol{\iota}_T'$. Also,

$$\boldsymbol{\omega}'\boldsymbol{\Omega}_e^{-1} = \frac{\sigma_u^2}{\sigma_v^2} b_i \left(\left(1 - \frac{T\sigma_u^2}{\sigma_1^2}\right)\boldsymbol{\iota}_T' \otimes \mathbf{B}_N\right)$$

where 1 is the $i$th column of $\boldsymbol{\iota}_T'$. Some manipulations and re-arrangements give

$$\boldsymbol{\omega}'\boldsymbol{\Omega}_e^{-1} = \frac{\sigma_u^2}{\sigma_v^2}\left(1 - \frac{T\sigma_u^2}{\sigma_1^2}\right) b_i \left(\boldsymbol{\iota}_T' \otimes \mathbf{B}_N\right)$$

$$\boldsymbol{\omega}'\boldsymbol{\Omega}_e^{-1} = \frac{\sigma_u^2}{\sigma_v^2}\left(1 - \frac{T\sigma_u^2}{\sigma_v^2 + T\sigma_u^2}\right) b_i \left(\boldsymbol{\iota}_T' \otimes \mathbf{B}_N\right)$$

$$\boldsymbol{\omega}'\boldsymbol{\Omega}_e^{-1} = \frac{\sigma_u^2}{\sigma_v^2}\frac{\sigma_v^2}{\sigma_1^2} b_i\left(\boldsymbol{\iota}_T' \otimes \mathbf{B}_N\right)$$

$$\boldsymbol{\omega}'\boldsymbol{\Omega}_e^{-1} = \frac{\sigma_u^2}{\sigma_1^2} b_i\left(\boldsymbol{\iota}_T' \otimes \mathbf{B}_N\right)$$

Because $b_i\left(\boldsymbol{\iota}_T' \otimes \mathbf{B}_N\right) = \left(1 \otimes b_i\right)\left(\boldsymbol{\iota}_T' \otimes \mathbf{B}_N\right) = \left(1 \cdot \boldsymbol{\iota}_T'\right) \otimes \left(b_i \cdot \mathbf{B}_N\right) = \left(\boldsymbol{\iota}_T' \otimes l_i'\right)$, where $l_i$ is the *i*th column of $\mathbf{I}_N$, I obtain the final expression for the Goldberger correction, i.e.

$$\boldsymbol{\omega}'\boldsymbol{\Omega}_e^{-1} = \frac{\sigma_u^2}{\sigma_1^2}\left(\boldsymbol{\iota}_T' \otimes l_i'\right)$$

Appendix 2 (Chapter 2)

**Ashton-under-Lyne**

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|---|---|---|---|---|---|---|---|---|
| 0 | -49.284 | | | | 31.4306 | 6.28549 | 6.28797 | 6.33378 |
| 1 | -14.5385 | 69.491 | 1 | 0.000 | .46336 | 2.06731 | 2.07226 | 2.16389 |
| 2 | -12.0416 | 4.9939* | 1 | 0.025 | .385508* | 1.88019* | 1.88761* | 2.02505* |
| 3 | -11.5573 | .9686 | 1 | 0.325 | .413791 | 1.94466 | 1.95455 | 2.1378 |
| 4 | -10.6447 | 1.8251 | 1 | 0.177 | .422882 | 1.95559 | 1.96795 | 2.19702 |

**Birmingham**

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|---|---|---|---|---|---|---|---|---|
| 0 | -34.5128 | | | | 4.95988 | 4.43909 | 4.44157 | 4.48738 |
| 1 | -11.0048 | 47.016 | 1 | 0.000 | .297909 | 1.6256 | 1.63054 | 1.72217 |
| 2 | -8.20479 | 5.6* | 1 | 0.018 | .238643* | 1.4006* | 1.40802* | 1.54546* |
| 3 | -7.65107 | 1.1075 | 1 | 0.293 | .253937 | 1.45638 | 1.46627 | 1.64953 |
| 4 | -7.60833 | .08548 | 1 | 0.770 | .289324 | 1.57604 | 1.5884 | 1.81747 |

**Blackburn**

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|---|---|---|---|---|---|---|---|---|
| 0 | -40.3937 | | | | 10.3449 | 5.17421 | 5.17668 | 5.2225 |
| 1 | -16.6183 | 47.551* | 1 | 0.000 | .60093* | 2.32728* | 2.33223* | 2.42386* |
| 2 | -16.4057 | .42511 | 1 | 0.514 | .665197 | 2.42572 | 2.43313 | 2.57058 |
| 3 | -16.2381 | .33519 | 1 | 0.563 | .742832 | 2.52977 | 2.53966 | 2.72291 |
| 4 | -16.1565 | .16323 | 1 | 0.686 | .842244 | 2.64456 | 2.65693 | 2.886 |

**Bolton**

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|---|---|---|---|---|---|---|---|---|
| 0 | -42.5233 | | | | 13.5001 | 5.44041 | 5.44289 | 5.4887 |
| 1 | -18.419 | 48.209* | 1 | 0.000 | .752627* | 2.55238* | 2.55732* | 2.64895* |
| 2 | -18.1808 | .47646 | 1 | 0.490 | .830448 | 2.6476 | 2.65502 | 2.79246 |
| 3 | -18.1807 | .0002 | 1 | 0.989 | .946991 | 2.77259 | 2.78248 | 2.96573 |
| 4 | -18.096 | .16941 | 1 | 0.681 | 1.07331 | 2.887 | 2.89936 | 3.12843 |

**Bradford**

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|---|---|---|---|---|---|---|---|---|
| 0 | -50.202 | | | | 35.2525 | 6.40025 | 6.40272 | 6.44854 |
| 1 | -36.2177 | 27.969 | 1 | 0.000 | 6.96324 | 4.77721 | 4.78215 | 4.87378 |
| 2 | -23.8537 | 24.728* | 1 | 0.000 | 1.68763* | 3.35671* | 3.36413* | 3.50157* |
| 3 | -23.7758 | .15588 | 1 | 0.693 | 1.90583 | 3.47197 | 3.48186 | 3.66512 |
| 4 | -23.717 | .11753 | 1 | 0.732 | 2.16707 | 3.58962 | 3.60199 | 3.83106 |

**Cardiff**

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|---|---|---|---|---|---|---|---|---|
| 0 | -56.1134 | | | | 73.8079 | 7.13918 | 7.14165 | 7.18747 |
| 1 | -15.7694 | 80.688 | 1 | 0.000 | .540428 | 2.22117 | 2.22611 | 2.31774 |
| 2 | -15.2966 | .94541 | 1 | 0.331 | .579085 | 2.28708 | 2.2945 | 2.43194 |
| 3 | -12.5961 | 5.401* | 1 | 0.020 | .471171* | 2.07452* | 2.08441* | 2.26766* |
| 4 | -12.5328 | .12665 | 1 | 0.722 | .535449 | 2.1916 | 2.20396 | 2.43303 |

**Edinburgh**

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|---|---|---|---|---|---|---|---|---|
| 0 | -66.6939 | | | | 277.001 | 8.46174 | 8.46421 | 8.51002 |
| 1 | -23.6721 | 86.044 | 1 | 0.000 | 1.45129 | 3.20902 | 3.21396 | 3.30559 |
| 2 | -23.1888 | .96675 | 1 | 0.325 | 1.55303 | 3.27359 | 3.28101 | 3.41845 |
| 3 | -19.9358 | 6.506* | 1 | 0.011 | 1.1793* | 2.99197* | 3.00186* | 3.18512* |
| 4 | -19.9333 | .00496 | 1 | 0.944 | 1.35041 | 3.11666 | 3.12902 | 3.35809 |

**Greenock**

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|---|---|---|---|---|---|---|---|---|
| 0 | -67.9922 | | | | 325.81 | 8.62403 | 8.6265 | 8.67231 |
| 1 | -31.3364 | 73.312 | 1 | 0.000 | 3.78288 | 4.16705 | 4.17199 | 4.26362 |
| 2 | -28.3749 | 5.923* | 1 | 0.015 | 2.96974 | 3.92186 | 3.92928 | 4.06672 |
| 3 | -26.9145 | 2.9208 | 1 | 0.087 | 2.82148* | 3.86431* | 3.8742* | 4.05746* |
| 4 | -26.1119 | 1.6052 | 1 | 0.205 | 2.92338 | 3.88899 | 3.90135 | 4.13042 |

Lag 3 selected.

**Halifax**

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|---|---|---|---|---|---|---|---|---|
| 0 | -53.6564 | | | | 54.29 | 6.83205 | 6.83453 | 6.88034 |
| 1 | -14.9763 | 77.36* | 1 | 0.000 | .489424* | 2.12204* | 2.12698* | 2.21861* |
| 2 | -14.8861 | .1804 | 1 | 0.671 | .550117 | 2.23576 | 2.24318 | 2.38062 |
| 3 | -14.6648 | .4425 | 1 | 0.506 | .610214 | 2.33311 | 2.343 | 2.52625 |
| 4 | -14.6604 | .00892 | 1 | 0.925 | .698583 | 2.45755 | 2.46991 | 2.69898 |

**Hull**

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|---|---|---|---|---|---|---|---|---|
| 0 | -37.2429 | | | | 6.97719 | 4.78036 | 4.78283 | 4.82865 |
| 1 | -24.3547 | 25.776 | 1 | 0.000 | 1.58056 | 3.29434 | 3.29929 | 3.39091 |
| 2 | -18.1497 | 12.41* | 1 | 0.000 | .827232* | 2.64372* | 2.65114* | 2.78858* |
| 3 | -17.8062 | .68699 | 1 | 0.407 | .903689 | 2.72578 | 2.73567 | 2.91893 |
| 4 | -17.7722 | .06816 | 1 | 0.794 | 1.03073 | 2.84652 | 2.85888 | 3.08796 |

**Leeds**

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|---|---|---|---|---|---|---|---|---|
| 0 | -55.6308 | | | | 69.4868 | 7.07885 | 7.08132 | 7.12714 |
| 1 | -13.7548 | 83.752* | 1 | 0.000 | .420122* | 1.96935* | 1.9743* | 2.06593* |
| 2 | -13.6712 | .16733 | 1 | 0.682 | .472606 | 2.0839 | 2.09131 | 2.22876 |
| 3 | -13.3378 | .66677 | 1 | 0.414 | .516939 | 2.16722 | 2.17711 | 2.36037 |
| 4 | -12.851 | .97363 | 1 | 0.324 | .557173 | 2.23137 | 2.24373 | 2.4728 |

**London**

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|---|---|---|---|---|---|---|---|---|
| 0 | -55.8974 | | | | 71.8411 | 7.11217 | 7.11464 | 7.16046 |
| 1 | -2.79597 | 106.2* | 1 | 0.000 | .106771* | .599496* | .604441* | .696069* |
| 2 | -2.56772 | .4565 | 1 | 0.499 | .117958 | .695965 | .703383 | .840825 |
| 3 | -2.17075 | .79394 | 1 | 0.373 | .128002 | .771343 | .781234 | .964491 |
| 4 | -1.10496 | 2.1316 | 1 | 0.144 | .128333 | .76312 | .775483 | 1.00455 |

**Manchester**

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|---|---|---|---|---|---|---|---|---|
| 0 | -36.8463 | | | | 6.63974 | 4.73079 | 4.73326 | 4.77907 |
| 1 | -22.4898 | 28.713* | 1 | 0.000 | 1.2519 | 3.06122 | 3.06617 | 3.1578 |
| 2 | -20.8973 | 3.185 | 1 | 0.074 | 1.16623* | 2.98716* | 2.99458* | 3.13202* |
| 3 | -20.7393 | .31603 | 1 | 0.574 | 1.3039 | 3.09241 | 3.1023 | 3.28556 |
| 4 | -19.7874 | 1.9038 | 1 | 0.168 | 1.326 | 3.09842 | 3.11078 | 3.33985 |

Lag 2 selected.

**Newcastle**

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|---|---|---|---|---|---|---|---|---|
| 0 | -64.0371 | | | | 198.725 | 8.12963 | 8.13211 | 8.17792 |
| 1 | -24.6297 | 78.815* | 1 | 0.000 | 1.63583* | 3.32871* | 3.33366* | 3.42529* |
| 2 | -24.4537 | .35208 | 1 | 0.553 | 1.81906 | 3.43171 | 3.43913 | 3.57657 |
| 3 | -24.4362 | .03485 | 1 | 0.852 | 2.06986 | 3.55453 | 3.56442 | 3.74768 |
| 4 | -23.241 | 2.3904 | 1 | 0.122 | 2.0419 | 3.53013 | 3.54249 | 3.77156 |

**Nottingham**

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|---|---|---|---|---|---|---|---|---|
| 0 | -52.056 | | | | 44.4464 | 6.632 | 6.63447 | 6.68029 |
| 1 | -13.6155 | 76.881* | 1 | 0.000 | .412867 | 1.95193 | 1.95688 | 2.04851 |
| 2 | -12.0445 | 3.142 | 1 | 0.076 | .385648* | 1.88056* | 1.88798* | 2.02542* |
| 3 | -11.6218 | .84525 | 1 | 0.358 | .417145 | 1.95273 | 1.96262 | 2.14588 |
| 4 | -10.8464 | 1.5508 | 1 | 0.213 | .433682 | 1.98081 | 1.99317 | 2.22224 |

Lag 2 selected.

**Sheffield**

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|---|---|---|---|---|---|---|---|---|
| 0 | -56.3503 | | | | 76.026 | 7.16879 | 7.17126 | 7.21708 |
| 1 | -14.5644 | 83.572* | 1 | 0.000 | .464861* | 2.07055* | 2.07549* | 2.16712* |
| 2 | -14.5225 | .08382 | 1 | 0.772 | .525671 | 2.19031 | 2.19773 | 2.33517 |
| 3 | -14.2123 | .62034 | 1 | 0.431 | .576653 | 2.27654 | 2.28643 | 2.46968 |
| 4 | -14.2116 | .00145 | 1 | 0.970 | .660469 | 2.40145 | 2.41381 | 2.64288 |

**Sunderland**

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|---|---|---|---|---|---|---|---|---|
| 0 | -61.523 | | | | 145.134 | 7.81537 | 7.81784 | 7.86366 |
| 1 | -25.2524 | 72.541* | 1 | 0.000 | 1.76826* | 3.40656* | 3.4115* | 3.50313* |
| 2 | -24.509 | 1.4869 | 1 | 0.223 | 1.83168 | 3.43862 | 3.44604 | 3.58348 |
| 3 | -24.462 | .09405 | 1 | 0.759 | 2.07652 | 3.55774 | 3.56764 | 3.75089 |
| 4 | -24.3944 | .1352 | 1 | 0.713 | 2.35854 | 3.67429 | 3.68666 | 3.91573 |

**Wigan**

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|---|---|---|---|---|---|---|---|---|
| 0 | -39.7237 | | | | 9.51379 | 5.09046 | 5.09293 | 5.13874 |
| 1 | -24.5383 | 30.371 | 1 | 0.000 | 1.61725 | 3.31729 | 3.32223 | 3.41386 |
| 2 | -21.7446 | 5.5874* | 1 | 0.018 | 1.29653* | 3.09308* | 3.1005* | 3.23794* |
| 3 | -21.4033 | .68272 | 1 | 0.409 | 1.41674 | 3.17541 | 3.1853 | 3.36856 |
| 4 | -21.2613 | .28392 | 1 | 0.594 | 1.59427 | 3.28266 | 3.29503 | 3.5241 |

**Wolverhampton**

| lag | LL | LR | df | p | FPE | AIC | HQIC | SBIC |
|---|---|---|---|---|---|---|---|---|
| 0 | -45.0385 | | | | 18.4875 | 5.75481 | 5.75728 | 5.80309 |
| 1 | -5.52381 | 79.029* | 1 | 0.000 | .150155* | .940476* | .945422* | 1.03705* |
| 2 | -5.25304 | .54153 | 1 | 0.462 | .165009 | 1.03163 | 1.03905 | 1.17649 |
| 3 | -4.93628 | .63354 | 1 | 0.426 | .180863 | 1.11703 | 1.12693 | 1.31018 |
| 4 | -4.39597 | 1.0806 | 1 | 0.299 | .19364 | 1.1745 | 1.18686 | 1.41593 |

## Appendix 3 (Chapter 3)

### Derivation of the Urban Economics Model

The starting point is a **Cobb-Douglas production function for the output *Q*** of the competitive constant-returns 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 intermediate market services inputs *I*** [29]. This means that the *I* sector is modelled as a 'continuum' of *x* varieties, each produced by a specialised firm with monopolistic / market power, with $i(z)$ representing the amount of type-*z* variety in the assumed composite intermediate market services sector.

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

In equilibrium, the CES sub-production function can be re-written as

$$I = x^{\mu} i(z)$$

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 market services, equal to $i(z)$, so that output $i(z)$ is constant across all varieties *z*. Parameter $\mu$ ($\mu > 1$) measures the equilibrium amount of **internal increasing returns to scale** that can be exploited by the individual firm, since an increase in *x* yields a more than proportionate increase in *I*.

[29] *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 firms producing highly-differentiated varieties, low entry and exit costs, and minimal strategic interaction, which is close to what is implied by monopolistic competition.

It also determines the <u>constant price elasticity of demand</u> ( $\varepsilon$ ), since from the constant elasticity demand function

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

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

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

and rearranging

$$\mu = \frac{\varepsilon}{\varepsilon - 1}$$

Parameter $\varepsilon$ ( $\varepsilon > 1$ ) denotes the **constant elasticity of substitution among varieties**. Thus $\mu$ also controls the <u>extent of differentiation in the *I* sector</u>, i.e. the preference of final goods and services producers for intermediate market services variety, and thus the <u>degree of monopolistic / market power available to *I* firms</u>, which is larger the more the *I* products are viewed by *M* firms as imperfectly substitutable.

This means that ***I*-sector internal scale economies $\mu$ arise from *M*'s love for variety and imperfect substitution $\varepsilon$**. As evident from $I = x^{\mu} i(z)$, the value of $I$ is not simply the sum of all *x* varieties, but reflects the added bonus that can be obtained from greater variety, unless $\mu$ approaches 1 and $\varepsilon$ goes to infinity in which case there is no benefit from variety.

To show that **variety *x* increases with density *L*** (i.e. effective employment per unit area), it is possible to write the number of intermediate market services firms/varieties, *x*, as labour efficiency units employed by the *I* sector in the unit area divided by *I*-sector effective labour per firm

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

where $\left[a \cdot i(z) + s\right]$ is the linear labour requirement function (labour is the only input, so production costs equal labour costs). Accordingly, the amount of labour efficiency units required to produce each *z*-type variety is given by a fixed component to start production, *s*, and a variable component which increases with output, $a \cdot i(z)$, 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 operate more productively by concentrating activity at one plant.

Behind this is the idea that 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 output level $i(z)$.

Replacing *I* with $x^{\mu} i(z)$ 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 $M = \beta \cdot L$ and for $x = \dfrac{(1-\beta)\cdot L}{a \cdot i(z) + s}$ I obtain

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

Collecting constants simplifies to

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

which shows that efficiency gains / cost advantages from **internal increasing returns to scale in the *I* sector translate into external scale economies** (productivity gains) **in the competitive, constant-returns final goods and services sector *M***.

# Appendix 4 (Chapter 3)

**Derivation of the Commuting-based Spatial Weights Matrix W**

The value allotted to cell (*i*, *j*) of $\mathbf{W}$ as given in section 3.4 (and section 5.2.3) is

$$W_{ij} = \exp(-\hat{\tau}_i d_{ij})$$

It is a function of the straight-line distance ($d_{ij}$) between areas *i* and *j* and of a **distance decay constant ($\hat{\tau}_i$) specific to area *i***. The latter controls, for *i*, the speed with which the relative weights of neighbouring areas decline as distance increases, thus **controlling the profile of spatially-lagged variables for *i***.

The procedure for selecting $\hat{\tau}_i$ is outlined below.

- <u>Calculate the *N* x *N* matrix of pairwise straight-line (Euclidean) distances</u> ($\mathbf{D}$). This requires LeSage's *distance* function from the Matlab spatial toolbox, which uses the *N* x 1 vectors of the x and y coordinates of the centroid of each UALAD (available from Ordnance Survey).
- <u>Calculate the *N* x 7 matrix of in-commuting percentages by UALAD and travel-to-work distance band</u>. The 1991 Census of Population provides data on commuting patterns (both inward and outward flows), from which I take the number of people travelling in for work from various UALADs of residence. Given the **D** matrix and seven distance bands (<2 km, 2 to 4 km, 5 to 9 km, 10 to 19 km, 20 to 29 km, 30 to 39 km, ≥40 km), it is possible to obtain for each (workplace) UALAD and band the observed proportions of in-commuting workers.
- <u>Estimate area-specific distance decay rates</u> ($\hat{\tau}_i$). The exponent $\hat{\tau}_i$ is chosen by iterating the function $\exp(-\tau_i d)$ through a range of values for $\tau_i$, starting from zero and progressively incrementing the value by 0.01, and using the upper limits of the six distance bands up to 39km to evaluate the function; the relative proportions for the distance bands are the simulated travel percentages. One hundred iterations are performed for each UALAD; the iterative process stops at the value of $\hat{\tau}_i$ which

maximises the fit to the area's commuting pattern, i.e. minimises the sum of squared differences between the observed and simulated travel percentages for *i*.

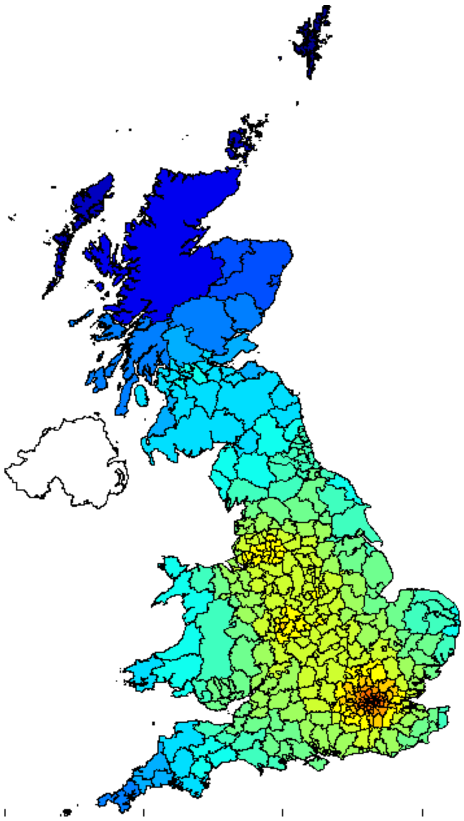
**Areas with a 'longer commuting distance profile' will have a smaller value of $\hat{\tau}_i$, since a smaller rate for *i* tends to attenuate the effect of distance, so that remote areas are more likely to affect *i*.**

**Areas with a 'shorter commuting distance profile' will have a larger value of $\hat{\tau}_i$; the larger the rate for *i*, the more rapidly the quantity $W_{ij}$ falls with increasing distance, so that remote areas are less or not important to *i*.**

Appendix 5 (Chapter 3)

**2008 Market Potential (MP) Relative to the Mean**

**Approximation (Harris, 1954)**



Note: Red: highest MP; Blue: lowest MP.

**Estimation of Krugman's non-linear, short-run function (Fingleton, 2006)**

Note: Black: highest MP; White: lowest MP.