Spatial dependence in unemployment in Great Britain

Master's Thesis

Alasdair Barrett - 2014

Submitted by Alasdair Barrett (383715) in July 2014 to fulfil the requirements of the MSc. in Urban, Port and Transport Economics at Erasmus University Rotterdam. Supervised by Jeroen van Haaren.

Acknowledgements

I would like to thank Jeroen van Haaren for his supervision and support of my project. His influence has substantially improved the quality of my thesis and given me many useful lessons for the future. I am also grateful to all of the lecturers on my course for providing me with the skills and inspiration to carry out this thesis project.

My girlfriend Steph, my family and all of my friends also deserve praise for helping me enjoy my year in the Netherlands and get the most out of my Master's programme. I am also thankful towards my colleagues at the Greater London Authority, particularly Matthew Waite, for giving me the opportunity to go on sabbatical and study for a year.

My final acknowledgements go to the developers of the STATA packages and open-source software that have made my thesis possible. These are;

- QGIS: Various developers
- GeoDa: Luc Anselin
- XSMLE: Federico Belotti, Gordon Hughes and Andrea Piano Mortari
- SPPACK: David Drukker, Hua Peng, Ingmar Prucha and Rafal Raciborski
- SG162: Maurizio Pisati

Executive Summary

In this study I find strong evidence that there is spatial dependence in unemployment in Great Britain. Spatial dependence is a situation where outcomes in a given area appear to depend on outcomes or other factors elsewhere and my study suggests that spatial dependencies or spatial spillovers are present in short-term, long-term and total unemployment in Great Britain.

One of the key results which has helped to establish this conclusion is that high levels of human capital in a region are associated with lower unemployment rates, not only in that region but in neighbouring regions also. This is a new result as although past research has examined unemployment in Great Britain for spatial dependence, my study is the first to attempt to examine the nature of these dependencies rather than just establish their presence.

The other insights from my research include that assuming that dependencies exist between regions up to 50km in distance from each other appears to give a better approximation of reality than when 25km, 100km, 200km or 300km are used as thresholds up to which dependencies are present. My study also suggests that the dependencies apparent in short-term, long-term and total unemployment are very similar.

These results are of clear societal relevance as they point towards two major policy suggestions. The first is that public authorities governing proximate regions, particularly those within 50km of each other, should collaborate on unemployment reduction policies so that beneficial spatial spillovers are accounted for in decision making. The second is that human capital improvements appear to be a beneficial policy option when dealing with unemployment, especially if combined with the collaboration also suggested.

My study was able to yield such relevant results by establishing a strong theoretical framework and designing a robust methodology. My theoretical framework made the case that as job-seekers carry out multi-regional job searches resulting in job-matches which, other things being equal, lower unemployment, there will be dependencies in unemployment between regions.

I also argued that the job-search behaviour of a job-seeker in a particular region is affected by their distance to this region, the attractiveness of working in this region and their perceived probabilities of getting a job in the region. This argument allowed me to establish a set of factors to include in my model by focussing on what could affect attractiveness and perceived success probabilities while remembering the role of distance.

The final element in my theoretical framework was the case I made for spatial dependencies to potentially differ for short-term and long-term unemployment. My argument for this adjustment was that job-search behaviour differs between the short- and long-term unemployed which may lead to differing dependencies in short- and long-term unemployment given that job-search helps explain these dependencies.

This theoretical framework helped me to assemble an appropriate dataset and design a suitable methodology. This methodology consisted of exploratory spatial analysis using the global and local versions of Moran's I which allowed me to demonstrate that unemployment in Great Britain is positively spatially correlated meaning proximate areas have similar unemployment rates. This correlation was apparent when considering the whole of Great Britain but appeared to be particularly driven by four large clusters three of which were high unemployment clusters in the North while the fourth was a Southern, low unemployment cluster.

Following this exploratory analysis I designed and carried out some econometric analysis. Baseline results using non-spatial panel models provided initial evidence regarding the links between human capital and unemployment but this was expanded on by the estimation of spatial models. A Spatial Durbin Model specification was used as this closely matched my theoretical model of unemployment and was statistically supported by the results of a number of tests.

The Spatial Durbin Models I estimated were able to provide evidence for the spatial dependencies I set out to examine which in turn allowed me to offer the conclusions and policy recommendations mentioned above. Full details can be found in the rest of my study which should be of relevance to anyone interested in unemployment in Great Britain or in spatial analysis in general.

Contents

Index of Figures

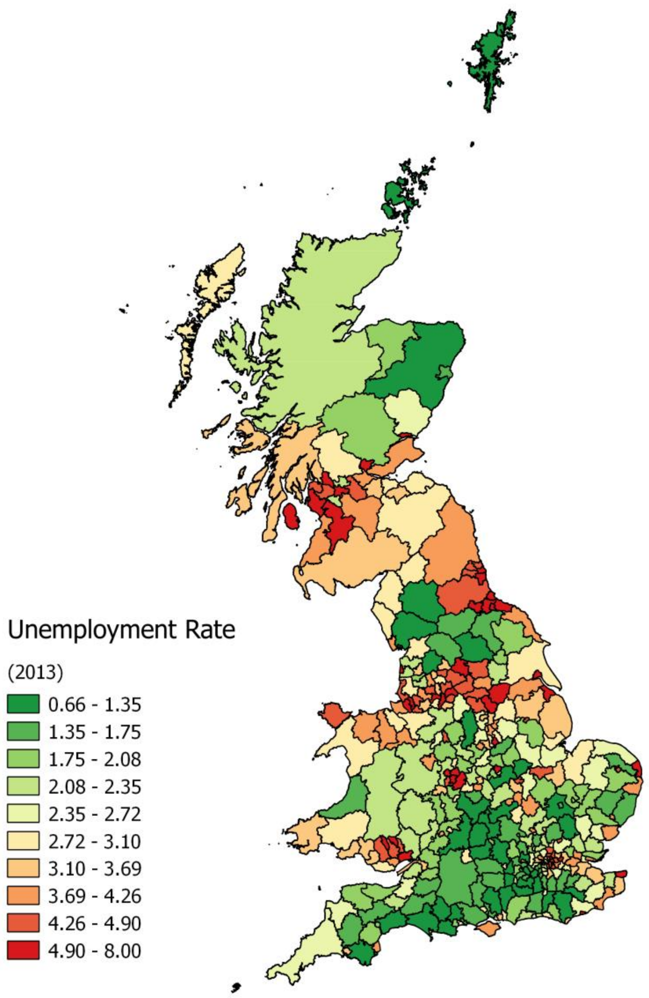
Index of Tables

1. Introduction

1.1. Background

The unemployment rate in Great Britain is actually lower than many of the other countries in Europe such as France, Italy or Spain (International Labour Organization, 2014), however, there are regions of Great Britain where unemployment is much higher than the national average. For example in 2013 there was a gap of 7.3 percentage points between the districts with the highest and lowest claimant count unemployment rates[1]. Many of these areas of labour market underperformance appear to be clustered together in certain locations, as shown in Figure 1, creating large swathes of the country where economic disadvantage is relatively common.

Figure 1: Total Unemployment Rate Deciles (2013)



Source: Author's analysis of Office for National Statistics (ONS) and Ordinance Survey (OS) data

The regional disparity in performance is even wider, in relative terms, when looking at long-term[2] unemployment as there are districts with long-term unemployment rates over three times as large as the average in Great Britain, whereas the worst performing region in terms of total unemployment has a rate just 2.3 times above the average. This means that short-term unemployment is more evenly distributed with the worst performing region only facing a short-term unemployment rate 2.0 times larger than Great Britain overall.

[1] Based on ONS data and using the resident population aged 16 to 64 as a denominator.
[2] Unemployment lasting over 52 weeks.

These regional variations in short-term, long-term and total unemployment are clearly important given the impacts unemployment can have at the individual and regional levels through, amongst other things, lowering individuals' quality of life and reducing regional output, respectively (Mankiw, 2008). This is particularly true given the apparent clustering in under and over-performing regions as this means there are large areas experiencing high unemployment.

Unsurprisingly, given these patterns, unemployment in the regions of Great Britain has been studied from many different perspectives. In the next section, I touch on this body of past research before explaining how this research has left some gaps in our understanding. The academic and societal relevance of filling these gaps with new analysis is also reinforced in the next section.

1.2. Problem Statement

Amongst recent papers aiming to describe or explain regional unemployment in Great Britain, some largely focus on national determinants (Nickell (1997), Blanchard (2006)), others do perform analysis at a regional level but assume regions are independent of each other (Webster (2005), Webster (2006), Theodore (2007), Brown and Sessions (1997), Little (2009)), while a third group analyses regional unemployment from a spatial perspective (Gilmartin and Korobilis (2012), Patacchini and Zenou (2007)).

Papers largely focussing on national factors include Nickell (1997) and Blanchard (2006). Both of these papers include the UK[3] within larger panels of countries and attempt to explain why experiences have differed between countries. Both papers offer useful findings regarding why the UK may have lower unemployment rates than some of its peers but unsurprisingly offer little explanation for the regional disparities and clustering that is observed in Great Britain.

Regional analysis is present in many papers focussing on unemployment in Great Britain but varies markedly in approach. Webster (2005), Webster (2006) and Theodore (2007) all take a non-econometric approach by analysing the relationship between total and long-term unemployment both for the country overall and for the constituent regions. These papers suggest that as total and long-term unemployment have a stable relationship over time, the reason some areas have higher unemployment rates is demand deficiency. They argue that if supply side issues, such as the unemployed becoming less employable, were present then the relationship between total and long-term unemployment would change, with each recession adding to the pool of long-term unemployed, few of whom would return to work when better economic conditions return.

[3] Great Britain is a subset of the UK

Econometric studies of unemployment with a regional dimension often consist of papers using individual level data to explain how regional factors affect employment outcomes alongside individual characteristics. Examples of this include Brown and Sessions (1997) and Little (2009). The former finds that living in certain regions of the UK (such as the North West or the North East) is associated with a higher chance of being unemployed, however, the paper doesn't try to explain this using data on the characteristics of these regions. The latter paper does, however, attempt to explain why some areas underperform in labour market terms, suggesting that demand shortfalls or mismatches between labour demand and supply may be to blame.

Turning to spatial analysis of regional unemployment in Great Britain, recent examples include Gilmartin and Korobilis (2012), and Patacchini and Zenou (2007). The former paper estimates a model of regional unemployment where the coefficients can vary between different clusters of regions. This approach incorporates spatial considerations by varying the coefficients between clusters, however, it does not actually include spatial dependence between regions.

Patacchini and Zenou (2007) does include such spatial dependence as it outlines a simple theoretical model where unemployment in an area partially depends on unemployment and other labour market factors in proximate areas. The empirical testing of this model finds support for this spatial dependence which helps explain the clustering of unemployment rates that is seen in Great Britain.

This body of past research can offer a number of useful insights, but where it is perhaps lacking, however, is a study that not only establishes the presence of spatial dependence but also aims to identify the spatial spillovers which contribute to dependence. Patacchini and Zenou (2007) does make some progress toward these aims but stops short of fully identifying and describing the spatial spillovers which influence the pattern of unemployment in Great Britain[4]. Another related gap is that much of the research into unemployment in Great Britain does not examine the more pronounced disparity in long-term unemployment relative to short-term or total unemployment, or examine the idea that spatial spillovers may be different for short- and long-term unemployment.

Addressing these gaps is important as unemployment has significant consequences at both the micro and macro levels. For individuals, being unemployed can result in lowered income and possibly self-esteem, while at the macro level, elevated unemployment levels entail a loss of economic output and raised welfare spending (Brown and Sessions, 1997).

[4] One of the reasons for this is that Patacchini and Zenou (2007) precedes LeSage and Pace (2009) which made several advances in the interpretation of spatial models. For more details on this, see section 4.2.3 later in this paper.

This problem of a lack of understanding concerning spatial spillovers affecting unemployment in Great Britain presents a clear motivation for carrying out this study. This is because demonstrating the geographic scale over which spillovers are apparent might strengthen the case for public authorities in proximate regions to collaborate on unemployment reduction schemes, and identifying the nature of the spillovers might point towards policy areas which should be focussed on. In order to maximise the contribution of this study toward solving the problem, it is important to set out clear research questions which can help steer the research design. I set out these research questions in the next section.

1.3. Research Questions

Given the problem statement I outlined in the previous section, the research questions and sub-questions I will attempt to answer are shown in Table 1. The answers to these questions will build up the evidence required to improve understanding of the spatial spillovers that affect unemployment in Great Britain. I have already addressed question I in the introduction, so the rest of the report will focus on the remaining questions.

Table 1: Research Questions

I.	What is the spatial pattern of unemployment in Great Britain?
II.	Why might unemployment vary between regions?
III.	Why might there be spatial dependence in unemployment?
IV.	Why might spatial dependence differ for short- and long-term unemployment?
V.	Is there spatial dependence in unemployment in Great Britain?
VI.	Is there spatial dependence in short-term unemployment in Great Britain?
VII.	Is there spatial dependence in long-term unemployment in Great Britain?
VIII.	**Are there differences in spatial dependence in short- and long-term unemployment in Great Britain?**

Question VIII is the main research question as the answer to this will establish whether there is spatial dependence in unemployment and whether these spatial spillovers differ for short- and long-term unemployment. Answering these questions will add to the body of knowledge concerning unemployment in Great Britain provided by existing spatial and non-spatial studies. These additions will relate to demonstrating the geographic extent and nature of spatial spillovers affecting short-term, long-term and total unemployment in Great Britain.

The answers to questions II, III and IV make up the theoretical framework required to answer the later research questions (V, VI and VII). I set out this theoretical framework in the next section in which I analyse a range of past research in order to explain why unemployment might vary between regions, why there might be spatial dependence in unemployment and why this spatial dependence may differ between different unemployment types.

Following the theoretical framework is a description of the data I used in the study and a section in which I set out and explain the various techniques used in the analysis. The results of my analysis are then presented and discussed in the next two sections which also include a summary of my study's limitations and the recommendations I have generated following my research. In the final section I conclude by summarising what my study has achieved overall.

2. Theoretical Framework

2.1. Introduction

As I explained in the introduction, the aim of the theoretical framework is to answer research questions II to IV and in doing so, set out an economic model of regional unemployment which can then be operationalised later in the paper using analytical techniques. In order to achieve this I first set out some basic concepts for non-spatial analysis of unemployment before describing how spatial spillovers may affect unemployment and how these spillovers may vary for short- and long-term unemployment.

In the following sections I do this by first describing traditional models of unemployment before moving on to a discussion of regional and spatial alternatives in which I outline the model I have used in this study. In the next section I consider the differences in short- and long-term unemployment and the implications of these differences for the model of unemployment I use in the study. The theoretical framework concludes with a summary of the economic model of regional unemployment which I have translated into an econometric model later in the study.

2.2. Traditional Models of Unemployment

Theoretical labour market analysis traditionally splits out unemployment into three components; structural unemployment, frictional unemployment and cyclical unemployment. To further complicate matters, the concept of a natural rate of unemployment, which is the rate of unemployment normally expected to prevail in an economy (consisting of both structural and frictional unemployment) is also used.

These categories offer helpful theoretical explanations for why unemployment exists. For example, structural unemployment exists when wages are above the market clearing rate, as this means labour supply will exceed labour demand. If wages were always able to adjust freely so that labour demand was equal to labour supply, then there would be no structural unemployment.

This, however, is unlikely to be the case as many countries (including Great Britain) have legally mandated minimum wages which are set above the market clearing rate, resulting in a certain level of structural unemployment. Structural unemployment may also result from wages being raised above the clearing rate by the collective bargaining actions of unions or by the choice of firms to pay efficiency wages (Mankiw, 2008).

Frictional unemployment exists when individuals spend time searching for a suitable job match. If people could instantly find and begin new jobs then there would be no frictional unemployment. Clearly this is also unlikely as individuals need to collect information on job openings, submit applications and perhaps undergo retraining. The level of frictional unemployment therefore depends on things such as the ease of finding and applying to job openings and the extent of average retraining requirements. Unemployment benefits also have an effect on frictional unemployment as they reduce the incentive individuals have to find a new job quickly (Mankiw, 2008).

Cyclical unemployment exists when the economy fluctuates from its trend growth path and firms respond by varying their hiring decisions. Without economic fluctuations there would be no cyclical unemployment and instead the level of unemployment would be determined by the amount of structural and frictional unemployment (Mankiw, 2008).

Despite the existence of these categories, measured unemployment, whether based on survey or administrative data, will include structural, frictional and cyclical unemployment together. This means that in practise empirical work tends to link theoretical factors relevant for a particular unemployment category to total unemployment rather than that category alone.

An example of this is Nickel (1997) which explains national unemployment differentials using variables linked to structural (e.g. union density), frictional (e.g. spending on active labour market interventions) and cyclical (e.g. inflation) unemployment. Using theories relating to all three components to explain observed unemployment need not be problematic, however, as the different unemployment components can be complementary in understanding the evolution of unemployment over time.

Blanchard (2006) makes this clear by surveying facts and theories which have been put forward to explain unemployment in the main European nations over the last 30 years. The paper's discussion is able to suggest an appealing and cohesive story using theories closely linked to all three of the unemployment components to explain why European unemployment increased markedly since the 70s but varied substantially between countries. This story is that economic fluctuations caused by the oil crisis pushed up observed unemployment, the extent of the increase was affected by variations in things such as collective bargaining in the different countries of Europe and the evolution from this point was influenced by changes to institutions such as unemployment benefits which also varied by country. This explanation draws on cyclical, structural and frictional unemployment highlighting the complementarities that exist.

In the same way that theories more closely linked to structural, frictional and cyclical unemployment can be complementary to each other, analysis of unemployment from a regional or spatial perspective can be complementary to the wealth of analysis at the cross-country level. This is because variations within countries can sometimes be as large as those between them, and because bordering regions in different countries often have very similar unemployment experiences despite their different institutions (Overman and Puga, 2002). Patterns such as this point towards the importance of regional factors as well as suggesting the potential existence of spatial spillovers which spatial research can help illuminate.

If analysis of unemployment did stop at the national level, there would be little explanation for such patterns. This would be problematic both as unemployment is often used as a key measure of regional performance and because large differences in unemployment rates between regions may be inefficient (Elhorst, 2003). It would also be inefficient from a research perspective as regional differences offer another source of unemployment variation that can be studied. In the next sections therefore, I go beyond traditional, national level, labour market analysis in order to explain why regional and spatial analysis can be beneficial for our understanding of unemployment.

2.3. Regional Models of Unemployment

Many models of regional unemployment have been suggested, but one of the most influential comes from Blanchard and Katz (1992). This model assumes that the different regions produce different bundles of goods under constant returns to scale and both firms and workers are perfectly mobile. The model is based around four relationships (short-run labour demand, wage setting, labour supply and long-run labour demand), but the reduced form simply relates unemployment to factors affecting labour demand, labour supply and wages.

This regional model can offer additional insights to national level analysis by helping explain why unemployment rates may vary regionally, something which is not clear from models focussing on national explanatory factors which vary little between regions (Elhorst, 2003). For example, in the steady state relationship, one way a region may have lower than average unemployment is if it experiences stronger than average labour demand, perhaps from increased demand for one of the products in the bundle of goods produced by the region or because regional infrastructure is particularly good (Blanchard and Katz, 1992).

Though useful, this steady state relationship explains unemployment for a region based only on factors in that region rather than including some kind of regional dependence. This is a problem for many studies of regional unemployment as most of the alternative models, such as single equation approaches, accounting identity approaches or other implicit models, result in the same reduced form equation (regional unemployment as a function of regional labour demand, labour supply and wage factors) as the Blanchard and Katz model (Elhorst, 2003).

Even where regions used for analysis do represent relatively distinct labour markets, the assumption of independence is unlikely to be viewed as realistic. More likely is a situation where proximate regions depend on each other in some way. Models allowing such spatial dependence, such as Vega and Elhorst (2013), Nistor (2009) and Patachini and Zenou (2007), can add to the insights from national and regional research as they can help identify or establish the presence of spatial spillovers that make regions dependent on each other. In the next section, I provide more details about these spatial models, and outline the spatial model of unemployment that was used in this study.

2.4. Spatial Models of Unemployment

As mentioned previously, a number of recent papers have examined unemployment from a spatial perspective. Vega and Elhorst (2013) is one example and uses the full Blanchard and Katz model as a starting point but adds an adjustment to allow for spatial dependence in the dependent and independent variables. The theoretical discussion in the paper focuses mainly on explaining the original Blanchard and Katz model but does also suggest that the adjustment for spatial dependence carried out may be justified theoretically as shocks in one region can also have impacts in neighbouring regions.

A less formal approach is taken in Nistor (2009) which doesn't directly set out an equation based model but instead starts from the reduced form of the Blanchard and Katz model and discusses a number of variables affecting labour demand, labour supply and wages, which are used to explain regional unemployment. The theoretical motivation for including spatial dependence is only given a short discussion but spatial externalities arising from behavioural, economic or political channels are cited as the justification for using a spatial approach. Nistor (2009) is fairly representative of the theoretical approach of a number of papers in this area such as; Aragon et al. (2003), Badinger and Url (2002) and Cracolici, Cuffaro, and Nijkamp (2007).

A different approach is taken in Patacchini and Zenou (2007) which specifies a simple two region theoretical model where unemployed workers can search for jobs in both their home region and the other region. This theoretical model suggests that unemployment in either of the areas depends on unemployment and labour market tightness in the other area as a result of the multi-area job search that unemployed individuals perform. This approach differs from other papers as the model used actually explains why there is spatial dependence in unemployment rates, rather than using (explicitly or implicitly) the Blanchard and Katz model and allowing spatial dependence as a relaxation of the model assumptions.

Through actually specifying the process through which spatial dependence in unemployment may arise, Patacchini and Zenou (2007) is better able to suggest how regions may depend on each other. Specifically, the paper proposes, and empirically verifies that, unemployment in a region is positively associated with unemployment in proximate regions and negatively associated with labour market tightness in nearby regions. Though the Patacchini and Zenou (2007) model has the advantage of being clear and explicit, it simplifies the situation found in most countries which in fact consist of multiple regions which have various proximities to each other rather than only two regions (which workers are assumed to search with equal intensity).

A paper which does address this real world complexity is Manning and Petrongolo (2011) which uses highly spatially detailed data to analyse the size of local labour markets using a job-search perspective. The paper models job-search behaviour as the result of a rational process where job seekers consider the costs and benefits of potential actions in their search. One of the elements of this model is that greater distance between job-seekers and vacancies they may apply to increases search costs so discourages search activity with respect to these vacancies.

Another useful element of the Manning and Petrongolo (2011) model comes from the other side of this cost and benefit determined search behaviour. Specifically, the model also describes how expected benefits, which relate both to the attractiveness of jobs within the potential search set and the chances of successfully gaining these jobs, encourage search activity. This gives a more realistic view of search activity in which individuals search multiple regions according to their distance to these regions, the attractiveness of jobs within them and the likelihood of obtaining one of these jobs.

The Patacchini and Zenou (2007) model is clearly simpler than this, but despite its simplifications, its clear and explicit nature make it a good framework to adapt in order to explain why there may be spatial dependence in regional unemployment. The key adaptations which will make the framework more realistic are incorporating the effects of distance, job attractiveness and chances of success which form part of the Manning and Petrongolo (2011) model. These adaptations of course relate to individual behaviours so their aggregate implications need to be considered if regional unemployment is to be explained.

Patacchini and Zenou (2007) moves from its discussion of individual job search behaviour to aggregate unemployment outcomes through the use of a matching function. The central idea of a matching function is that it describes how job vacancies are filled by job seekers whose search behaviour has previously been described. Matches can therefore be seen to lower unemployment, other things being equal, meaning they can be included in a model of the unemployment rate. Patacchini and Zenou (2007) therefore goes through three simple steps to explain spatial dependence in unemployment. The first is to underline that job seekers search in multiple regions, the second builds upon this by stating that this multi-regional search behaviour affects successful job matches and the third step is to explain that successful job matches, other things being equal, reduce unemployment.

Using these same three steps, but adding the small adaptations mentioned previously, results in the following model which I used in this study. Firstly, job seekers search multiple regions but their search in these regions depends on the distance to them, the attractiveness of working in them and the perceived likelihood of getting a job in them (following Manning and Petrongolo (2011)). Secondly, this multi-regional search behaviour results in a certain number of successful job matches which thirdly, other things being equal, reduce unemployment. This is summarised below in Figure 2.

Figure 2: Adapted Patacchini and Zenou (2007) Model

$$Unemployment_i = f\left(Job\ Matches_{i,i}\ , Job\ Matches_{i,j}\right) \qquad \text{for all j}$$

$$Job\ Matches_{i,j} = f(Distance_{ij}\ , Attractiveness_j, Perceived\ Success\ Probabilities_j\ )$$

Where 'i' and 'j' represent regions so $Job\ Matches_{i,j}$ represents matches of residents in region i with jobs in region j (where j can equal i)

As this model focuses on job-search and matches to explain spatial spillovers in unemployment, it most closely links to frictional unemployment out of the various categories outlined in my discussion of traditional models of unemployment (section 2.2). Given that measured unemployment includes not only frictional but also structural and cyclical unemployment, the methodology (which I describe in section four) makes adjustments for this to better capture the spatial spillovers which are of relevance for my research questions.

It is also important to note that in this model job matches can, of course, be within a job-seeker's home region (i) or in any other region (j). These matches depend on distance, attractiveness and on perceived success probabilities as these determine job-seekers' search effort. In this way my model includes spatial spillovers as distances between a home-region and other regions as well as other factors in these non-home regions affect unemployment in the home-region.

An obvious question is therefore, how to capture these other factors of attractiveness and perceived success probabilities. I will outline this later in the theoretical framework, as in the next section I address the important issue of why this model may vary if short- and long-term unemployment are considered separately.

2.5. Differences Between Short- and Long-term Unemployment

In the previous section I outlined how job-search behaviour can be used to explain spatial dependence in unemployment rates. In this section I explain that, to the extent to which there are differences in job-search between the short- and long-term unemployed, this may mean there are differences in spatial dependence seen in short- and long-term unemployment.

Potential differences in job-search between the short- and long-term unemployed are part of the debate over why long-term unemployed individuals tend to have inferior job prospects to individuals with shorter unemployment durations. The two main arguments in this debate are the existence of unobserved heterogeneity or the idea that unemployment has a scarring effect (Brown and Sessions, 1997).

Unobserved heterogeneity is a situation where the least employable of the unemployed are different in some way which means they have longer periods of unemployment. Alternatively it may be that being unemployed has a scarring effect with individuals' skills and motivation perhaps eroding away, meaning that, other things being equal, a candidate that has been unemployed longer would be less desirable for a firm. Clearly these arguments can apply to more than just job-search, but nevertheless job-search is relevant for both lines of argument. For example, if the long-term unemployed are in fact different in some way then this may manifest in their job search behaviour, similarly their job search behaviour may change over time if unemployment scars them in some way.

One way in which the long-term unemployed may be different is their location of residence which is an important determinant of job accessibility and job-search behaviour. For example, Détang-Dessendre and Gaigné (2009) finds that improved access to jobs (through proximity) results in shorter periods of unemployment. This is consistent with the theoretical model in Wasmer and Zenou (2006) within which individuals are faced with distance related search costs and can choose their search effort. In this scenario, unemployed people with good access to jobs face lower commuting costs to interviews, this means they search more intensely and have shorter spells of unemployment. Conversely, the higher search costs of those with poor access to jobs mean they search less intensively and have longer spells of unemployment.

Interestingly, Smith and Zenou (2003) actually shows that selecting a residence with poor access to jobs and performing low intensity job searches can be an optimal scenario for unemployed individuals. This is because the short run benefits of such a choice set, lower housing costs and greater housing consumption, may outweigh the longer term benefits of the alternative, higher re-employment probability due to increased job access. There is some empirical support for this idea that individuals in remote areas may search for jobs less intensely, for example see Patacchini and Zenou (2005).

In addition to these ways in which unobserved heterogeneity may mean search behaviour varies between the short- and long-term unemployed, there are also arguments related to the scarring view. For example, it could be argued that as individuals face longer out of employment their resources may become depleted leading them to lower their reservation wage and accept more distant job offers. Ahn, De La Luca and Ugidos (1999), however, actually finds there is no difference in willingness to relocate for work between individuals with different unemployment durations, but those who have remained unemployed so long that their benefits run out do appear more willing to relocate.

Another way unemployment could affect job search behaviour is through lowering the motivation of job seekers, meaning they spend less time performing searches. Krueger and Mueller (2010) analyses this but finds that as long as unemployed individuals are still receiving unemployment insurance the amount of time they spend searching for jobs stays at similar levels regardless of their duration of unemployment.

Overall, therefore, there is some evidence that job search may vary between the short- and long-term unemployed. Though not conclusive, there appears to be some support for the idea that the long-term unemployed search less intensely due to having inferior job access which means they face greater search costs. As mentioned, the idea that job search may vary between the short- and long-term unemployed suggests that the dependence job search creates between regions may differ for these two groups. Given the specific differences that have been outlined, it is expected that spatial dependence will be less apparent in long-term unemployment.

This means that the relationships outlined previously in Figure 2 may be different depending on whether short- or long-term unemployment is being considered. In order to fully explore this possibility later in the paper, it is necessary to outline the specific factors which will be included in my model of unemployment as potential influences on job-search behaviour and therefore matches and unemployment. In the next section I describe these factors using insights from past research.

2.6. Fundamental Factors

The adapted version of the Patacchini and Zenou (2007) model which I have used in this paper, explains that job seekers search multiple regions depending on their distance to them and various factors relating to these regions. This multi-regional search behaviour results in a certain number of successful job matches which, other things being equal, reduce unemployment. In order to operationalise this model, it is necessary to set out which factors are relevant and so should be included in the model.

I aim to do this in the following passages by describing how factors such as demographics, participation rates and wages are of relevance. These factors will of course be included in the model multiple times given the multi-regional job search which forms the basis of this model. This means that, for example, unemployment in district 'i' will be affected by demographics in region 'i' and demographics in other regions. The directions of these direct and indirect effects need not always be the same and this is something which I have also discussed in the sections below.

2.6.1. Demographics

Age-structure can influence the unemployment rate of an area through its impact on search behaviour as young workers are more likely to change jobs in search of a good match than their older peers who have already had a chance to go through this process (Brown & Sessions, 1997). This suggests areas with many young workers would be expected to have higher unemployment rates as, at a given time, more people will be between jobs searching for an appropriate match.

Aragon et al. (2003) finds empirical support for this theory in a study of the Midi-Pyrenees region in France as does Lottman (2012) for Germany.

The share of older workers in the labour force may also be important as older workers may have greater experience of going through the job search process having done this previously in their careers. This may mean their perceived chances of success are higher so they search more intensely and, therefore that regions with a greater share of older workers in the labour force experience lower unemployment.

The effects of having a labour force that is relatively young or old on search activity and unemployment in neighbouring regions, referred to as the indirect effect, are perhaps less clear cut. It could, however, be argued that individuals searching a non-home region may perceive their chances of securing a job to be higher, other things being equal, if the labour force in that region is younger and less experienced. Alternatively, the opposite may be true where a labour market is full of workers who are more experienced on average.

2.6.2. Participation Rates

Participation rates also have implications for search behaviour which mean they can affect unemployment. Elhorst (2003) suggests that higher participation rates are associated with lower unemployment as they encourage job growth, something which the paper suggests is supported by most of the empirical evidence on the matter. This job growth would influence search behaviour by increasing labour market tightness (as there would be more vacancies relative to job seekers) and therefore perceived chances of success.

Alternatively, it could be argued that higher participation rates could reduce labour market tightness by increasing the number of job seekers relative to vacancies. Fleisher and Rhodes (1976) finds some empirical evidence pointing in this direction. Given that participation rates could potentially influence job search, and therefore matches and unemployment in two directions, it is difficult to state exactly what direct effect would be expected though a negative relationship seems more common empirically.

The same is true for the indirect effect as if higher participation rates stimulate employment growth this could be beneficial for unemployment in neighbouring regions. However, if they merely mean there are more job-seekers per vacancy, the opposite could be true. Again, given that the empirical evidence tends to suggest higher participation rates are associated with lower unemployment, this would be expected for the indirect effect also.

2.6.3. Wages

There are a number of ways in which wages can influence search behaviour and therefore unemployment and many views have been suggested. One argument is that higher wages increase job attractiveness and therefore encourage search activity. This would mean that places with higher wages tended to have lower unemployment rates, much like the wage curve view in Blanchflower and Oswald (2005) which suggests that in high unemployment areas, workers will not press for high wages as they have a weaker bargaining position.

Alternatively it may be that higher wages suggest there are tougher requirements for successful candidates to meet which reduces perceived success probabilities and therefore search activity. This would mean places with higher wages tend to have higher unemployment rates much like the Harris-Todaro view in which individuals will only agree to work in high unemployment areas if they have a high wage to compensate for the greater unemployment spells they expect to experience (Blanchflower and Oswald (2005)).

Blanchflower and Oswald (2005) surveys a number of empirical studies on this issue which they report have tended to support a negative relationship, where higher wages are associated with lower unemployment, for the majority of countries that have been analysed. Elhorst (2003) also suggests that studies examining the link between wages and unemployment have tended to find a negative relationship, however, this paper also makes the point that many studies fail to use a measure of real wages when this is what actually influences decisions.

The indirect effects of higher wages could also act in both directions for the same theoretical reasons. As above, however, a negative relationship would be expected given that this is more common empirically.

2.6.4. Employment Growth

Higher rates of employment growth are expected to have a negative relationship with unemployment as with more jobs, perceived chances of success are higher, search intensity is greater and there are more job matches. The empirical evidence surveyed in Elhorst (2003) on this relationship does appear to broadly support the theoretical negative relationship as do more recent studies such as Lottman (2012) for Germany or Gilmartin and Korobilis (2012) for the UK. The indirect effect of employment growth on unemployment would also be expected to be negative, as new jobs in a given region will also be beneficial for the perceived chances of success of residents in neighbouring regions.

Although an argument could be made that employment growth may itself be associated with high unemployment regions (as firms want to take advantage of the slack labour market), this does not appear to be the case. For example, Overman and Puga (2002) suggest that employment growth has actually contributed to widening disparities in unemployment in European regions as it has tended to be higher in areas already performing well. This provides additional support to the expectation of negative direct and indirect effects.

2.6.5. Industry Mix

A region's industry mix could influence search behaviour and therefore unemployment through altering perceived chances of success as for example, regions with a large share of their employment in a sector that is declining nationally may be viewed as offering inferior success prospects. There is, however, limited empirical support for this with many studies finding that industry mix doesn't explain much of the regional variation in unemployment (Elhorst, 2003), perhaps as the same industry can experience different fortunes in different regions (Martin, 1997).

Another channel of impact relating to industry mix comes from changes in the industry mix which may be expected to increase unemployment. This may be because changes to industrial structure can mean job-seekers' skills are less suited to vacancies on offer so their perceived chances of success and search intensity drop. Elhorst (2003) suggests that measures of the extent of industrial change do tend to be positively related to regional unemployment.

A final way in which industry mix may affect unemployment is through industrial diversity offering opportunities for redeployment. Regions with a greater variety of industries are likely to offer more opportunities for workers who lose their jobs in a declining industry to find another industry that is a suitable fit, meaning they may perceive their chances of success to be greater. Elhorst (2003) suggests that most of the empirical studies that have looked at this issue have found supporting evidence for this claim.

The industry mix, changes to the industry mix and industrial diversity would all be expected to have indirect effects in the same direction as their direct effects. This is because their effects on perceived chances of success would also be relevant and of the same direction for residents in neighbouring regions.

2.6.6. Housing

Housing status can affect search behaviour and therefore unemployment in a number of ways, for example, residents in public housing may be reluctant to move region in search of a job. This may be because receiving a place in public housing is effectively a subsidy (Brown and Sessions, 1997) and the prospect of losing this subsidy by moving for a new job may make public housing residents' job searches more distance constrained. The same may also be true of owner-occupiers given the transaction costs associated with home purchase. This limited mobility of owner-occupiers relative to renters is supported by Muellbauer and Murphy (1991) which also confirms those in public housing are even less mobile than owner-occupiers.

Having a tenure profile concentrated on less mobile tenure types could have a beneficial effect on unemployment in neighbouring regions. This may be the case as residents in neighbouring regions would face less competition for jobs in their home region from non-home region residents who were more distance constrained. This would increase their perceived chances of success and encourage search activity suggesting the indirect effect of having large proportions of social housing tenants or owner occupiers would be negative.

Housing is also of relevance as house prices may reflect the attractiveness of working in a particular area or could reflect the labour market prospects from living in such an area. For example, houses in locations with good access to jobs or in low unemployment regions may be expected to have higher prices. This would mean higher house prices were associated with lower unemployment.

If higher house prices are reflective of superior labour market prospects in an area, then the effect of being situated near high house price regions would be expected to be negative on unemployment rates. This is because the good labour market prospects which may have contributed to high housing prices will also be accessible to residents in nearby regions and so would be expected to lower unemployment in these regions.

2.6.7. Benefits

The real value of unemployment benefits can affect unemployment rates through increasing the reservation wage of job seekers by reducing the attractiveness of getting a new job to them (Elhorst, 2003). As with public housing recipients, benefits recipients may therefore be more distance sensitive given the subsidies they are receiving.

Even if such benefits are set nationally, if this is done in nominal terms then they may have an impact at regional level if different living costs mean their real value varies. This channel is highlighted in Little (2009) which explains that the real value of some UK benefits is highest in areas with low employment levels.

Proximity to regions where the real value of benefits is high might have a similar negative effect to being proximate to regions with a high proportion of social housing tenants or owner occupiers. This is because generous benefits may also make individuals more distance sensitive.

2.6.8. Human Capital

Individuals with higher levels of human capital are expected to be more consistently in demand by employers and when out of work are expected to conduct more efficient job searches (Elhorst, 2003). This means they would have higher perceived chances of success and that areas with a high proportion of individuals with higher levels of human capital are expected to have lower unemployment rates.

Human capital may also have an effect on search activity if it stimulates job growth and therefore increases labour market tightness. This could be the case if greater human capital improves productivity leading to economic expansion and job creation. This creation of jobs would also suggest regions with higher human capital experience lower unemployment, something which does appear to be the case according to the empirical evidence surveyed in Elhorst (2003).

The effect of being proximate to regions with high human capital would tend to reduce unemployment if the productivity benefits of a concentration of skilled workers do stimulate job growth. Greater levels of human capital in a neighbouring area could also act as a signal that the area includes a large proportion of firms in knowledge intensive industries. This could also be positive for the attractiveness of jobs in this area, encouraging search activity and therefore also pointing to higher levels of human capital being associated with lower unemployment in proximate regions.

An argument could be made that proximity to regions with higher human capital might increase unemployment if residents perceive their chances of success to be lower as a result of competition from the more skilled residents in the proximate region. This effect could, however, be masked if the employment growth and job attractiveness channels have a stronger influence, as is expected.

2.6.9. Amenities

The possible theoretical links between amenities and job search behaviour are similar to those mentioned with respect to wages. On one hand it could be argued that having good amenities increases the attractiveness of working in an area, encouraging job search and lowering unemployment.

Alternatively it could be argued that the presence of favourable amenities may make job-seekers anticipate greater competition for roles in an area, discouraging their job search activities. Elhorst (2003) and Aragon et al. (2003) both argue that better amenities would be associated with higher unemployment, however the empirical evidence is far from definitive.

By the time of the review of empirical evidence in Elhorst (2003), only two studies had analysed this effect, both finding insignificant effects. A more recent study, Lottman (2012), does find significant effects for amenity variables, however, these effects are not in the expected direction, perhaps because the amenity proxies are somewhat tenuous given they include the public debt ratio and the numbers of business registrations.

The indirect effects of amenities could also be either positive or negative for unemployment given the contrasting arguments that have been made. The argument that amenities make an area more attractive encouraging job search and lowering unemployment is in line with the evidence on wages so it is logical to expect the direct and indirect effects of amenities to be negative for unemployment.

2.6.10. Unemployment

Unlike the previous factors, it is clearly not appropriate to consider the direct effects of unemployment, however, unemployment in neighbouring regions may influence a given region of interest, and indeed studies such as Patacchini and Zenou (2007) make this case. These effects would be expected to be positive as high unemployment in a neighbouring region suggests a loose labour market where chances of success are low. As explained previously this would dissuade search activity and increase unemployment.

2.7. Summary of Theoretical Framework

In the previous sections I have outlined how spatial analysis of unemployment can be complementary to analysis done from a national or regional perspective through illuminating spatial spillovers between regions. I have also set out the economic model my paper will use in investigating spatial dependence in unemployment in Great Britain by answering research questions II, III and IV

In essence my answers to these questions are that unemployment can vary between regions as regions may produce different bundles of goods and experience different labour market developments such as demand increases or decreases. The unemployment rates of these regions may be dependent on factors in other regions as job-seekers search for jobs and achieve job-matches in multiple regions.

This multi-regional job search process depends on distances between regions as well as the attractiveness and potential for successfully finding a job within them. As job search behaviour affects job matches which lower unemployment, other things being equal, then these same factors are of relevance for regional unemployment. In a situation of spatial dependence these factors have direct effects (on the region they relate to) and indirect effects (on neighbouring regions), both of which are of interest.

My answers allow hypotheses to be stated regarding the remaining research questions. For research questions V, VI and VII, following the discussion in the theoretical framework I expect that spatial dependence will be present given the links multi-regional job search creates between different areas. For the main research question the theoretical framework I have established again points to a likely answer through explaining that there is evidence the long-term unemployed search less intensely and have inferior job access to the short-term unemployed which would suggest spatial dependence is less apparent in long-term unemployment than short-term unemployment.

The next section outlines the data that I have used in order to answer the remaining research questions and assess whether these hypotheses should be rejected. In this section I first discuss the dependent variable before outlining details of the independent variables I used.

3. Data

In the following sections I outline the data that I used to answer my key research question of whether there are differences in spatial dependence in short- and long-term unemployment in Great Britain. In the first sub-section I focus on the dependent variable while in the second I cover the independent variables.

3.1. Dependent Variable

The focus of my study is spatial dependence in unemployment and therefore the dependent variable I used is the unemployment rate. According to the International Labour Organisation, anyone who is out of work but is searching for employment, and available to start, is unemployed while people doing any amount of paid work, people temporarily away from work (e.g. on holiday) or people doing unpaid work for a family firm are employed (Office for National Statistics, 2013).

The employed and unemployed make up the economically active and the unemployment rate is therefore the number of unemployed people as a proportion of those who are economically active. In my study the measure used will differ slightly from these ILO guidelines as the number of unemployed will be captured using the number of people claiming unemployment benefits, known as the claimant count.

I made this decision for two major reasons. Firstly, the claimant count measure is administrative data and so gives an accurate picture of claimants even when small geographical units are used for analysis, an important consideration given my paper uses spatial analysis. This would not be the case if the ILO definition of the unemployment rate was used as, for Great Britain, published ILO unemployment figures are derived from a sample survey (the Labour Force Survey) and so are less reliable for small areas.

The second reason for using the claimant count measure is that claimants can be readily segmented according to the duration of their claim allowing for the analysis of short- and long-term unemployment that the paper aims to do. This is also not possible using publically available figures which adhere to the ILO definitions.

A second difference the data I have used in this paper have with the ILO definition is that the denominator used is the total working age population, rather than the active population. I made this choice as figures on the number of active labour market participants in an area would also rely on sample surveys which are less accurate for small geographical areas.

I sourced the desired measure from Nomis, an online portal for labour market statistics from the Office of National Statistics (ONS). I used the data for 379 of the 380 districts in Great Britain from 2004 to 2013 in this study[5]. The district level was chosen as all the key independent variables are available at this level of disaggregation, something which is not true of alternative units, whether smaller, such as Lower Layer Super Output Areas, or larger, such as Travel-to-Work Areas.

The measure of total unemployment I used in this study captures all claimants, while those claiming for up to one year are classed as short-term unemployed and those claiming for longer than one year are long-term unemployed. This threshold of one year is now used as standard according to Webster (2005). Summary statistics for the three measures are shown in Table 2 and it is also informative to look at histograms of short- and long-term unemployment as these highlight that long-term unemployment appears to have a longer tail.

Table 2: Summary Statistics for Unemployment Measures

	Mean	Standard Deviation (S.D.)	Min	Max	Count
Total Unemployment Rate	2.62	1.37	0.43	8.79	3790
Short-Term Unemployment Rate	2.19	1.07	0.40	6.92	3790
Long-Term Unemployment Rate	0.43	0.39	0.00	3.28	3790

[5] The Scilly Isles were excluded as many of the independent variables were not available for this district.

Figure 3: Histogram of short-term unemployment

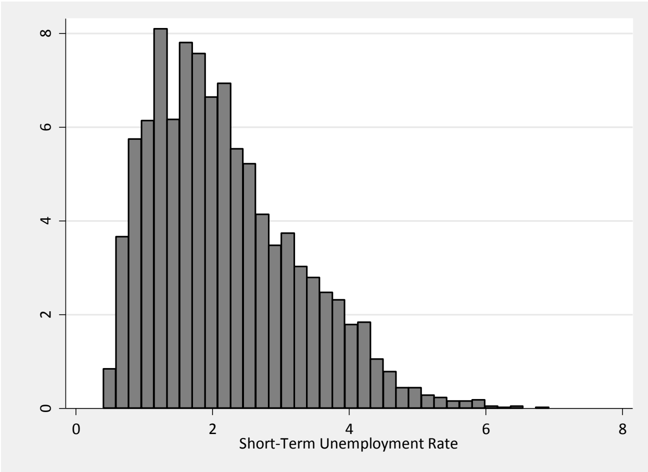
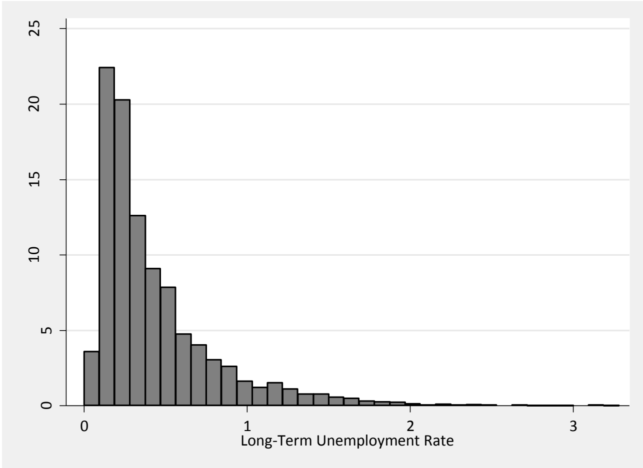


Figure 4: Histogram of long-term unemployment



In this section I explained that the dependent variable will deviate slightly from the ILO definition of unemployment. This deviation ensures that data are accurate at a low geographic level and can disaggregate short- and long-term unemployment. The next section outlines the independent variables which have been used in the study.

3.2. Independent Variables

In the theoretical framework section I explained how the multi-regional job search process carried out by job seekers can cause spatial dependence in unemployment. This multi-regional search process was stated to depend on distances between regions as well as the attractiveness and potential for successfully finding a job within them. Factors affecting attractiveness and success potential were also outlined and now I will describe the exact measures used to represent these factors as well as the data used to establish distances between regions. I have included descriptive statistics for the key explanatory variables in Table 16 and correlations between these variables in Table 17, both of which can be found in the appendix.

3.2.1. Demographics

The first factor to be included in the model is demographics and this was captured through the use of two variables. The first measures the proportion of the working age population that are aged 16 to 24 and the second measures the proportion of the working age population that are aged 50 to 64. On average 17% of the working age population are aged 16 to 24 and 30% are aged 50 to 64.These data are from the ONS Annual Population Survey (APS) and are similar to measures used in Aragon et al. (2003).

3.2.2. Participation Rates

Also sourced from the APS is the measure of participation rates used in the study. This measure gives the proportion of the working age population that are economically active which is 78% on average. Participation rates are also used in Aragon et al (2003).

3.2.3. Wages

A measure of median gross weekly earnings is used to represent wages and is taken from the ONS Annual Survey of Hours and Earnings (ASHE). On average, median gross weekly earnings are around £370 but the variable has been log-transformed for inclusion in the model. I did this in order to ease interpretation and to reduce the correlation between this variable and the proportion of working age residents with a degree or above. It was not possible to use a measure of real wages as suggested in Elhorst (2003) because there are no regional price indices that could be used for deflation.

3.2.4. Employment Growth

Two alternative measures of employment growth were collected for the dataset. The first is also from ASHE and is likely to be somewhat imprecise as the focus of this survey is estimating earnings not employment. The alternative measure is from the APS and is expected to be more accurate, however, unlike the ASHE based measure this is only available for nine of the 10 years of the study period. The ASHE based measure captures the change in the number of jobs in a district while the APS measure captures the changes in the number of people employed in a district. Similar variables are also used in Gilmartin and Korobilis (2012).

3.2.5. Industry Mix

The dataset I assembled contains various measures of the industrial composition of the districts, all sourced from the APS. These measures include employment shares in each of the nine broad sectors of the economy[6] and a measure of industrial concentration, the share of a district's employment that is in its three largest sectors. On average, 66% of the labour force works in the three largest sectors in a district. Measures of industrial concentration such as this are fairly common in studies of regional unemployment according to Elhorst (2003). No measure of the annual extent of industrial change was available for use.

[6] Agriculture and Fishing; Energy and Water; Manufacturing; Construction; Distribution and Hospitality; Transport and Communications; Banking, Finance and Insurance; Public Administration, Education and Health; Other Services

3.2.6. Housing

Data on the housing tenure profile at regional levels was not available for all the required districts, however, information on median house prices at the district level was available. This median price is around £175,000 on average, however, unfortunately this data is somewhat problematic as the measurement process differs between that used by the Land Registry in England and Wales[7] and the Registers of Scotland[8]. This meant that I treated the consolidated series with caution in my analysis. Nevertheless, house prices are also relevant for job-search behaviour having been used in studies such as Patacchini and Zenou (2006).

An alternative categorical transformation of the house price measure is also included in my dataset and splits the house price distribution into thirds each year. I carried out this transformation as the original measure of house prices was strongly correlated with the proportion of working age residents with a degree or above. More details on the correlations can be found in Table 17 in the appendix.

3.2.7. Benefits

A measure of the proportion of the resident, working-age population that are claiming any kind of benefit, except unemployment benefits, has been used as no information on the real value of benefits by district was available. This measure averages 11% and includes working age residents claiming benefits because they are disabled, on low income, have caring responsibilities or are widowed. The data series has been sourced from the Work and Pensions Longitudinal Study and is therefore administrative data, not a sample survey.

The measure of benefits is quite strongly correlated with the measure of participation rates and with the measure capturing the share of residents without any qualifications. More details on the correlations can be found in Table 17 in the appendix.

3.2.8. Human Capital

The dataset I assembled includes two measures of human capital. The first is the proportion of working age residents with a degree (or equivalent) and above. The second measure is the proportion of working age residents with no qualifications. On average 21% of the working age population have a degree or above while 12% have no qualifications. Both of these measures are based on APS data and are similar to measures used in Badinger and Url (2002).

[7] Who exclude properties sold at prices above £20m from the data collection
[8] Who exclude properties sold at prices above £1m from the data collection

3.2.9. Amenities

The regional level of amenities could not be added to the dataset as no appropriate measure was available.

3.2.10. Unemployment

Unemployment is included in the model but only with respect to indirect effects (the effects unemployment in neighbouring regions has on unemployment in a home region). The same data used for the dependent variable is also used for this.

3.2.11. Distances

Distances between the districts in Great Britain are a key factor affecting unemployment as they have an influence on job-search activity. The distances are those between the central points of the districts and I calculated these in kilometres using a shape-file of district borders from the Office of National Statistics and Ordinance Survey.

In the previous sections I outlined the dataset assembled for my analysis, details of the methodology I used for the analysis can be found in the next section. These details include an explanation of the exploratory analysis and of the econometric modelling which follows it.

4. Methodology

4.1. Exploratory Analysis

In this section I describe the exploratory analysis that I performed for this study. The aim of this exploratory analysis was to provide some initial evidence regarding the main research question and sub-questions V to VII which concern the existence of spatial dependence in unemployment. Generating this initial evidence was also useful in providing statistical justification to go along with the theoretical justification for the use of spatial models in the econometric analysis. My exploratory analysis used a number of spatial analysis techniques and the description of these techniques aims to provide sufficient detail so that results are easy to understand even to those new to the area of spatial analysis.

The first technique I used was the calculation of a measure of global spatial autocorrelation. Measures of global spatial autocorrelation are designed to examined the extent to which there are spatial patterns in data, something which is expected in unemployment given the map (Figure 1) in the introduction showed considerable clustering. Such patterns, if found, can provide preliminary evidence of spatial dependence in the analysed variables. A leading measure of this type, used in studies such as Badinger and Url (2002), Patacchini and Zenou (2007) and Rae (2012), is Moran's I.

Much like a standard correlation coefficient, the Moran's I ranges between -1 and 1, with values close to 1 indicating positive spatial autocorrelation, where similar values are found close to each other, and values close to -1 indicating that places close to each other tend to have different values. Values near to 0 would therefore be expected if there was no spatial dependence (Rae, 2012). This makes the Moran's I easy to interpret and is why I decided to use this measure instead of alternatives such as Geary's c and Getis and Ord's G.

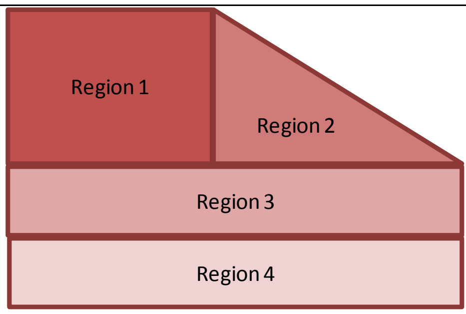
To calculate Moran's I statistics requires not only data with a spatial dimension but also information about the layout and location of the spatial units the data relate to. This is because the calculation of Moran's I requires what is called a spatial weight matrix as can be seen in Table 3. Such a matrix describes how the different geographical units relate to each other.

Table 3: Moran's I

Equation $I = \frac{n}{S_0} \frac{\sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2}$
Variables I Moran's I n number of observations x_i variable to be tested w_{ij} element in the spatial weight matrix $S_0 = \sum_i \sum_j w_{ij}$ sum of elements in the weight matrix i, j - districts *Based on (Fischer & Wang, 2011)*

These descriptions can be based around contiguity (whether or not the geographical units border each other), distance or even other factors such as bilateral migration flows for example. Whatever the basis, the end result is a square grid with the number of rows and columns equal to the number of spatial units. In each cell within the grid is a measure of the relation between the spatial units that have been allocated to the row and column in which the cell is situated. Cells on the main diagonal of the grid are set to zero (as it is not necessary to capture the association a spatial unit has with itself) and often the values within the grid are normalised so that each row sums to one as in the example shown in Figure 5.

Figure 5: Example Spatial Weight Matrix



Geographical arrangement of example regions

	Region 1	Region 2	Region 3	Region 4
Region 1	0.00	0.50	0.50	0.00
Region 2	0.50	0.00	0.50	0.00
Region 3	0.33	0.33	0.00	0.33
Region 4	0.00	0.00	1.00	0.00

Contiguity-based, row-normalised spatial weight matrix

As the spatial weight matrix is a key component in the calculation of the Moran's I, different choices regarding how to make this matrix affect the Moran's I produced. As such, it is relatively common to calculate multiple Moran's I values using different weight matrices. Vega and Elhorst (2013) is a good example of this approach as it uses ten different types of weight matrix while Patacchini and Zenou (2007) also uses several variants.

Another issue with using the Moran's I is inference. This is important for interpreting the Moran's I values obtained from any calculation as without some appreciation of whether a value is significantly different from 0 it would be difficult to use such a value as preliminary evidence for any spatial dependence. The two approaches used to address this issue are random permutation procedures and use of approximate sampling distributions. The latter was the approach I used and essentially involves using an approximate sampling distribution for the Moran's I to calculate a z-score which asymptotically follows a normal distribution and can therefore be compared against a standard normal table to assess significance (Fisher and Wang, 2011; StataCorp, 2001).

Therefore my overall approach to examining global spatial autocorrelation was to calculate the Moran's I using multiple spatial weight matrices and assess the significance and magnitude of these measures. As my theoretical framework argued that regions are dependent on each other due to multi-regional job searches carried out by job-seekers and that these searches are affected by the distance between job seekers and jobs, I used distance-based spatial weight matrices (but cross checked the results using a contiguity-based matrix also). Linking the choice of spatial weight matrix to theory is important, as Elhorst (2010) suggests that this is seldom done in spatial econometric studies and Corrado and Fingleton (2012) suggests the selection of a weight matrix is often the least theoretically backed part of a spatial econometric paper.

$$I_i = (x_i - \bar{x}) \sum_{j \in J_i}$$

The specific matrices I used for calculating Moran's I values contained inverse-distance weights which were row normalised. I used distance thresholds of 25km, 50km, 100km, 200km and 300km meaning that weights only had non-zero values where the distances between districts were below these thresholds. I also generated a row normalised Queen contiguity[9] matrix which was used as a cross check.

Through calculating multiple Moran's I values I was able to make comparisons between total, short-term and long-term unemployment in terms of the significance and magnitude of the statistics using weight matrices with different distance thresholds. This provided some preliminary evidence for the main research question and sub-questions V to VII by indicating whether there was global spatial autocorrelation in total, short-term and long-term unemployment.

In addition to the calculation of global measures of spatial autocorrelation, my preliminary analysis also involved calculating measures of local spatial autocorrelation. The measure I used was the Local Moran's I, which is very similar to the global Moran's I but can give an indication of which areas are contributing to the overall Moran's I and therefore the presence of global spatial autocorrelation (Rae, 2012).

Calculations of the Local Moran's I can be plotted on maps which are able to show spatial clusters (areas of positive spatial autocorrelation) and spatial outliers (areas of negative local spatial autocorrelation) providing a rigorous demonstration of whether apparent clusters do actually demonstrate significant local spatial autocorrelation (Anselin, 2005). The equation for the local Moran's I is given in Table 4. As was the case with global spatial autocorrelation, there are alternative measures (such as the local versions of Geary's c and Getis and Ord's G), but I selected the local Moran's I both for its ease of interpretation and its similarity to the global measure used. This analysis of local measures of spatial autocorrelation also provided some preliminary evidence of relevance to my main research question and sub-questions V to VII by highlighting areas where spatial autocorrelation in unemployment appeared particularly apparent.

Table 4: Local Moran's I

Equation

$$I_i = (x_i - \bar{x}) \sum_{j \in J_i}^{n} w_{ij}(x_j - \bar{x})^2$$

[9] Queen contiguity means that districts sharing either borders or vertices are classed as contiguous

Variables
I_i Local Moran's I
n number of observations
x_i variable to be tested
w_{ij} element in the spatial weight matrix
J_i neighbourhood set of district i
$\bar{x}$ mean, within the neighbourhood set
i, j - districts
Based on (Fischer & Wang, 2011)

The preliminary analysis not only provided some initial evidence regarding the main research question but also justified my use of spatial econometric models in the later analysis by indicating the presence of spatial autocorrelation. My econometric analysis then investigated this autocorrelation further by estimating spatial spillovers between regions rather than just testing the extent to which proximate regions have similar experiences. The next section explains what my econometric analysis consisted of.

In this section I have outlined the exploratory analysis techniques which I used to provide some initial evidence regarding the research questions and to justify the spatial modelling carried out in the study. This exploratory analysis involved using spatial weight matrices to calculate measures of global and local spatial autocorrelation in order to determine whether spatial patterns which may be striking visually are significant statistically. In the next section I outline the econometric techniques which will be used to generate additional evidence to help answer my research questions.

4.2. Econometric Analysis

In this section I outline the econometric analysis which I performed after the more simple, univariate exploratory analysis. I estimated a number of models to provide additional information to help answer my research questions and the details of these models will be described below. As with my description of the exploratory analysis, in this section I aim to provide sufficient information for readers new to the area of spatial econometrics.

4.2.1. General Approach

The dataset I assembled included repeated observations on the same spatial units over time so all of the models estimated were panel models. The first models I estimated were non-spatial and were used to generate some baseline results regarding the associations between the explanatory variables and unemployment. I followed this by estimating of spatial models which were of more direct relevance to my main research question.

The use of spatial models was designed to provide evidence on any spatial dependence in unemployment through the estimation of spatial spillovers, but was also pursued for statistical considerations as ignoring spatial dependence in an econometric model may lead to inefficient and possibly biased estimates (Chasco (2013), Lottman (2012), Foote (2007)). Starting with a non-spatial model before moving on to spatial alternatives is a common approach, taken in papers such as Fuhr and Sunde (2002), Longhi and Nijkamp (2007) and Lottman (2012). In order to re-confirm the appropriateness of using spatial models, I tested the residuals from the non-spatial models for global spatial correlation by calculating Moran's I values.

For all of the models, I used fixed-effects variants in order to control for time invariant unobserved characteristics. One key unobservable that should remain largely fixed over time is the set of environmental amenities each district provides and as amenities are one of the factors which could not be explicitly included in the model, using the fixed effects versions was helpful. The choice of fixed effects was also verified using a Hausman test to make sure using fixed-effects did not entail an unnecessary loss of efficiency.

I also included time fixed effects in the models. This was in order to control for general economic circumstances affecting the whole of Great Britain during the ten years of the panel. The use of both area and time fixed effects should secure a closer match between my econometric and economic models of unemployment as these fixed-effects may capture structural and cyclical components of observed unemployment, leaving the estimated parameters to describe the effects included in my theoretical framework which was most closely related to frictional unemployment.

4.2.2. Selection of Spatial Models

There are three common types of spatial model (though more exist) and these are set out in Table 5. The first model, the Spatial Durbin Model (SDM), includes spatially lagged independent variables and a spatial lag of the dependent variable. These spatially lagged variables are the result of multiplying a vector of observations for a particular variable with a spatial weight matrix. So for example, the spatially lagged dependent variable in this case would, for each district and year, give the average unemployment rate of proximate[10] districts.

[10] Where proximity is defined according to the rules used to generate the spatial weight matrix and the average is distance weighted.

A spatially lagged variable is also included in the second model type, the Spatial Autoregressive Model (SAR). Unlike the SDM, however, the SAR only includes a spatially lagged dependent variable, not any spatially lagged independent variables. The third model does not include any spatially lagged variables at all and is the Spatial Error Model (SEM) which instead includes a spatially correlated error component.

The various motivations for using spatial models such as time dependence, omitted variables, spatial heterogeneity and externalities can suggest one model may be more appropriate than another (LeSage and Pace, 2009). In this case, my theoretical framework does suggests the SDM might be most appropriate as it includes spatially lagged dependent and independent variables. This is very similar to my theoretical framework in which I stated that unemployment in a particular region may be affected by unemployment in other regions (justifying the inclusion of a spatially lagged dependent variable) and other factors in these different regions such as demographics (justifying the inclusion of spatially lagged independent variables).

It is not certain, however, that the SDM represents the true data generating process but in such a situation of model uncertainty, LeSage and Pace (2009) recommends starting from the SDM anyway. This is because the SDM subsumes the two other major model types and has the advantage of producing unbiased estimates even if the true economic process follows the SAR or SEM variants (Elhorst, 2010).

Table 5: Spatial Panel Model Equations

General Equation $y_{it} = \alpha + \rho W y_t + X_{it}\beta + W X_t \theta + \gamma_t + \mu_i + \varepsilon_{it} + \lambda W \nu_t$
Variables y_{it} dependent variable α constant W spatial weight matrix X_{it} independent variables γ_t time fixed effects μ_i district fixed effects ε_{it} error ν_t spatially correlated error component i - districts t - time periods

<u>Model definitions</u>

Spatial Durbin Model; $\rho \neq 0 \quad \theta \neq 0 \quad \lambda = 0$

Spatial Autoregressive Model; $\rho \neq 0 \quad \theta = 0 \quad \lambda = 0$

Spatial Error Model; $\rho = 0 \quad \theta = 0 \quad \lambda \neq 0$

Based on; Belotti, Hughes and Mortari (2013), Hughes (2012) and Lundin (2013)

My approach, therefore, was to use the SDM as the first spatial model estimated before testing to ensure it was the most appropriate selection. This is a similar approach to that used in Aragon et al (2003), Lottman (2012) and Nistor (2009).

The first test required aims to establish that the spatially lagged dependent variable (ρ) is significant but the spatially lagged independent variables (θ) are jointly insignificant as this would suggest the SAR is more appropriate. Alternatively, a second test aims to assess if $\theta = -\rho\beta$ as if this were the case then the SDM model becomes the SEM model (Elhorst (2014)) which would therefore be more appropriate. If neither of these restrictions are supported then the SDM is preferred to either alternative.

4.2.3. Estimation of Spatial Models

My estimation of the selected spatial models was via maximum likelihood using the STATA package XSMLE made by Beloti, Hughes and Mortari (2013). This package not only estimates the model coefficients but also produces summary measures of the average direct, indirect and total effects of the explanatory variables. These direct, indirect and total effect estimates follow the procedures in LeSage and Pace (2009).

The production of summary measures for the direct, indirect and total effect of a variable in a spatial model[11] is useful as the coefficients in such models are not as simply interpreted as those in Ordinary Least Squares models. This is because a change in an independent variable in a certain location can affect the dependent variables in other locations in the model. There is further complexity from feedback effects where a change in one region affects its neighbours which then effects their neighbours, one of which is the original region (LeSage and Pace (2009)).

The summary measures, therefore, help deal with this complexity by providing more easily understandable information on all the effects contained within the parameter estimates. Presenting

[11] these effects are only present in spatial models including spatial lags such as the SDM and SAR, not the SEM

these measures is not only helpful but important too, as their directions can vary from those of the parameter estimates as can their statistical significance.

Of the three summary measures, the average direct effect captures the effect of a unit change in an explanatory variable for one region, on the dependent variable of that region, averaged over all the regions. The average indirect effect is the effect of a unit change in an explanatory variable in all regions other than one, on the dependent variable in that region, averaged over all regions. Finally, the average total effect captures the effect of a unit change in an explanatory variable for all regions, on the dependent variable in one region, averaged over all the regions (LeSage and Pace (2009).

Though the package XSMLE used in the analysis is very helpful in calculating these measures, it has one significant constraint. Specifically, the package requires a perfectly balanced panel which had implications for the variables and districts that could be included in the model. This is particularly apparent when looking at the descriptive statistics in Table 16 as some of the variables do have missing values.

4.2.4. Selection of Spatial Weight Matrices

Selecting a particular spatial model type and estimation procedure are not the only important decisions in spatial modelling, much like the calculation of Moran's I, spatial weight matrices must also be chosen when doing spatial econometrics. This of course means that the choice of matrix can affect the results and means it is sensible to test multiple matrices and select the matrix which appears to provide the best model fit.

Doing this has become a relatively common feature of spatial studies because the fact that the weight matrix is not estimated but specified in advance is viewed as a weakness of spatial econometric models (Elhorst, 2010). The same set of spatial weight matrices that I used in the calculations of global spatial autocorrelation were the ones I tested in the econometric models. This was to provide consistency and to continue to reflect the role of distance which was highlighted in the theoretical framework. The selection of the matrix which provides the best fit was done using a comparison of Akaike Information Criteria (AIC) values, similar to the approach in Vega and Elhorst (2013). The results I present later relate to models using the matrices with the lowest AIC values but these results were checked for robustness using a comparison with results I obtained when using a Queen contiguity matrix.

As in Klinger and Rothe (2012), I estimated separate models for total, short-term and long-term unemployment. This allowed the results to be compared to help provide an answer to the main research question. My comparisons used three channels. Firstly, I compared the type of model

(SDM, SAR or SEM) which appeared to give the best fit for each unemployment type. Secondly, I compared the spatial weight matrix which gave the best model fit for each type and my third and final comparison was of the sign and significance of the direct, indirect and total effects.

4.2.5. Alternative Approaches Rejected

The alternative econometric approaches which I considered but decided to reject were dynamic spatial panel models, mixture panel data models, multi-level modelling or spatial filtering. Dynamic spatial panel models, where variables are lagged in time as well as space, are used in Vega and Elhorst (2013), but this is a relatively new area of research, for example Elhorst (2009) explains that the literature in this area is still developing, and this made it less suitable for use in my current study.

Gilmartin and Korobilis (2012) uses mixture panel data models, an approach where estimated coefficients can vary between data-determined clusters. As the clusters in this method are data determined, this approach lacks the theoretical underpinning achieved in my paper. Mixture panel data models also have the disadvantage of not being able to shed any light on spatial spillovers.

Multi-level modelling, another alternative, is often utilised when dealing with regional data as it allows errors to be calculated which account for the clustered nature of the data. An issue with this approach is that allocating the geographical units of analysis to larger clusters (in Great Britain this would mean clustering the districts within Government Office Regions) does not deal with correlations that cross these cluster borders (Foote, 2007). This is clearly a disadvantge of the multi-level modelling approach as is the fact that a multi-level model with no spatial terms would not provide any information about spatial spillovers.

Lack of information about spatial spillovers was also my primary concern with the final econometric technique considered; spatial filtering. This is used in Badinger and Url (2002) and involves removing spatial correlation from the dependent and independent variables using a filter before estimating a model using the filtered variables.

4.2.6. Summary

In this section I have outlined the econometric methodology used in my study. This methodology involved estimating non-spatial models to provide baselines results and then spatial models to provide results with greater relevance to the research questions.

I chose, for theoretical and statistical considerations, a fixed-effects SDM specification as the starting point for the spatial modelling and then tested whether this choice was appropriate. I estimated these SDM models using a recent STATA package which also provided summary measures of the direct, indirect and total effects of the variables included. My choice of weight matrices was done using AIC values. The results I obtained from this approach, which was chosen after consideration of various alternatives, are set out in the next section.

5. Results

5.1. Exploratory Analysis

My exploratory analysis involved calculating measures of spatial autocorrelation using a variety of different spatial weight matrices. The results for the Moran's I measure of global spatial autocorrelation are shown below in Table 6. Of the six spatial weight matrices I used, five are row-normalised, inverse-distance matrices with threshold distances as stated. The final matrix is a row-normalised, Queen contiguity matrix.

Table 6: Moran's I Measures of Global Spatial Autocorrelation

	Total Unemployment		Short Term Unemployment		Long Term Unemployment	
Spatial Weight Matrix	2013	2004	2013	2004	2013	2004
25km	0.43	0.42	0.41	0.41	0.45	0.42
50km	0.36	0.33	0.36	0.34	0.35	0.28
100km	0.24	0.22	0.24	0.23	0.24	0.20
200km	0.18	0.18	0.19	0.19	0.17	0.15
300km	0.15	0.15	0.15	0.16	0.14	0.13
Contiguity	0.40	0.47	0.39	0.46	0.42	0.49

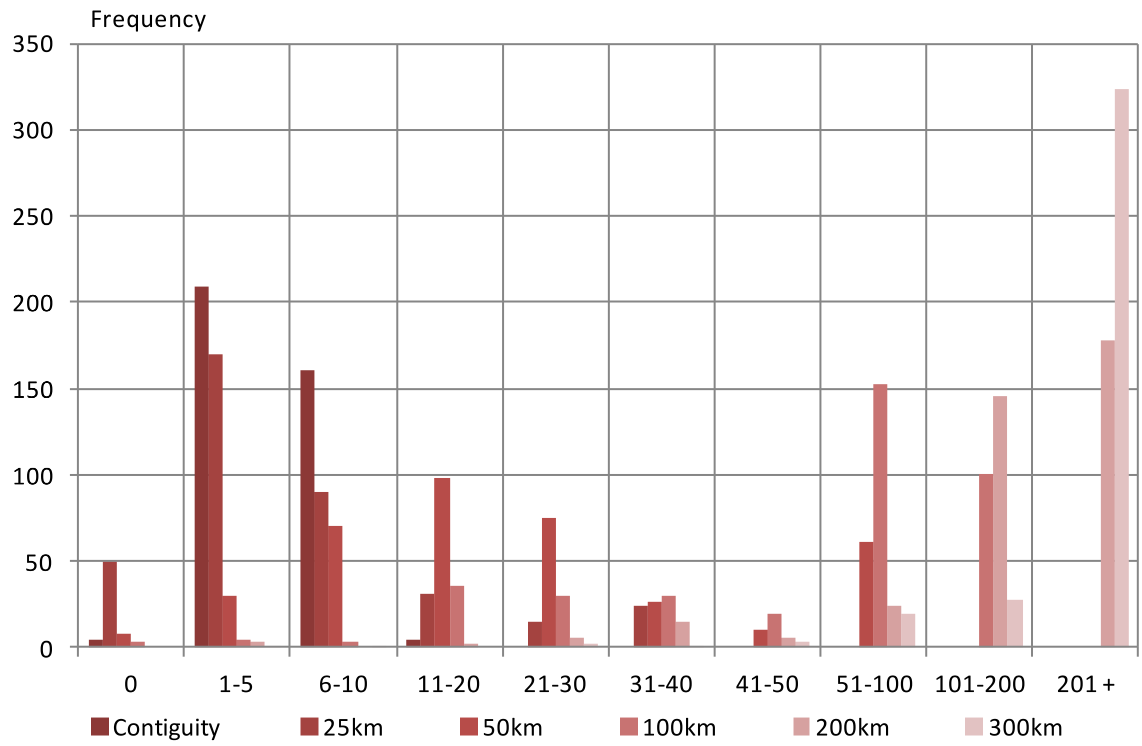
All measures are significant at the 0.1% level

All of these measures of spatial autocorrelation are highly significant and show positive spatial autocorrelation. This means that regions close to each other tend to have similar rates of total, short-term and long-term unemployment as was seen in the map in section 1 (Figure 1). For each unemployment type, the strength of spatial autocorrelation declines as the distance threshold used in the spatial weight matrix increases, for example the Moran's I for total unemployment declines from 0.43 to 0.15 when the distance threshold of the spatial weight matrix increases from 25km to 300km. This means that as the set of regions considered to be neighbours expands from those up to 25km away to those up to 300km away, the correlation of unemployment rates between these neighbours lessens.

The spatial autocorrelation measures resulting from the use of a contiguity matrix are similar in magnitude to those from the matrix with a 25km distance threshold. This is not too surprising given the similarity in the connectivity structure of the two matrices, which can be seen in Figure 6.

This figure shows, for example, that the majority of districts have between one and five neighbours in the 25km and contiguity spatial weight matrices, unlike some of the matrices with greater distance thresholds which result in neighbourhood sets with far greater numbers of members.

Figure 6: Connectivity Structure of Spatial Weight Matrices



The differences in the Moran's I values between the unemployment types are relatively minor as are those between the years of 2004 and 2013. This means that the main findings from this first piece of exploratory spatial analysis are that there is positive spatial autocorrelation in all three categories of unemployment which appears stronger at shorter distances but does not appear to differ dramatically over time or between the unemployment types.

Spatial autocorrelation is also present in some of the explanatory variables that were used in the models[12]. Moran's I values for the main measures used are shown in Table 7 and have been calculated using 2013 data and the spatial weight matrix with the 50km threshold. There is positive spatial autocorrelation in one of the measures of demography and in the measures of participation rates, human capital, wages and house prices.

[12] As explained later, some of the variables in the dataset were not included in the final models

Table 7: Spatial Autocorrelation in Independent Variables

	Moran's I	P Value
Proportion of Working Age Population Aged 50 or Above	0.21	0.00
Proportion of Working Age Population Aged 16 to 24	0.00	0.78
Working Age Participation Rate	0.21	0.00
Proportion of Working Age Population with a Degree or Above	0.47	0.00
Share of Employment in Largest Three Sectors	0.03	0.10
Median Gross Weekly Earnings	0.42	0.00
Median House Price	0.66	0.00

The second element of my exploratory analysis aimed to see if particular regions are responsible for the observed global spatial autocorrelation in the dependent variable. I achieved this by calculating values of the Local Moran's I and plotting a map showing the type of spatial autocorrelation present in all of the locations for which the Local Moran's I is significantly different from zero.

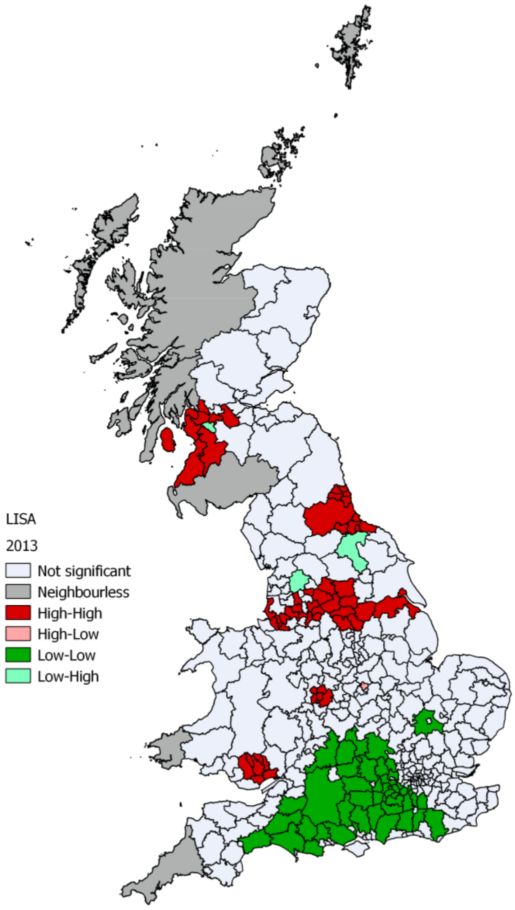
Such a map can be seen in Figure 7, which plots areas with significant Local Moran's I values in 2013 (the equivalent figure for 2004 can be seen in Figure 8, in the appendix). There are six categories on the map; districts with no neighbours (when using the 50km distance-threshold spatial weight matrix), districts where the Local Moran's I is insignificant and four categories of district where the Local Moran's I is significant.

These four categories refer to the type of cluster the region is in, and these clusters can be 'high-high' or 'low-low', indicating positive local spatial autocorrelation or 'high-low' or 'low-high', indicating negative local spatial autocorrelation. The first word refers to the district itself while the second refers to its neighbours, as defined by the spatial weight matrix. This means a district marked 'high-high' has a higher than average unemployment rate as do its neighbours.

Looking at the map, there are four large clusters and a few smaller areas of clustering. Given that the global Moran's I demonstrated positive spatial autocorrelation, it is not surprising that these four major clusters demonstrate the same. The three Northern clusters (from most Northern to most Southern) cover Glasgow and the surrounding areas, the Tyneside region and a large area of the midlands from Lincolnshire in the East to Liverpool in the West. These Northern clusters are all 'high-high' clusters with above average unemployment. The major Southern cluster goes from Sussex at its Easternmost to Devon in the West and is a 'low-low' cluster. The clusters highlighted in this map are unsurprisingly similar to the areas of clustering which are visually apparent in Figure 1.

Overall, my exploratory analysis clearly demonstrates that there is positive spatial autocorrelation in unemployment in Great Britain. This autocorrelation appears strongest when the distance threshold for districts to be considered neighbours is relatively short at 25km or 50km, or if contiguity is used to define neighbouring relationships.

Figure 7: Total Unemployment Cluster Map (2013)



Source: Author's analysis of Office for National Statistics (ONS) and Ordinance Survey (OS) data. 50km spatial weight matrix and 5% significance level used. LISA stands for local index of spatial autocorrelation

There are four major clusters and a few smaller clusters that seem to particularly contribute to the global autocorrelation. Of the major, clusters, three are in the North of Great Britain and are associated with elevated unemployment rates, while the only Southern cluster is associated with low unemployment. This points toward the North-South divide mentioned in past research such as Brown and Sessions (1997).

In the next section I build on these initial findings by estimating a number of spatial and non-spatial econometric models. The results of these models provide evidence which help to answer my main research question and the preceding sub-questions.

5.2. Econometric Analysis

My exploratory analysis has established that there is positive spatial autocorrelation in unemployment in Great Britain. In this section I outline the results of my econometric analysis which involved estimating models allowing for spatial dependence between regions in order to gain insights into the spatial spillovers which may contribute to this observed correlation. Before performing any spatial modelling, however, I first estimated non-spatial models to provide some results for comparison.

I used the same variables in both the spatial and non-spatial models in order to facilitate comparison, however, not all of the measures I outlined in section 3.2 were used in the final models. I excluded the benefits measure as it captured the prevalence of benefits receipt rather than the real value of benefits, which is what my theoretical framework highlighted as important. It also had the disadvantage of being highly correlated with the participation rate, a variable which did closely link to the theoretical framework. In addition, I did not include either of the employment growth measures as using the measure deemed to be higher quality by the ONS (the APS-based measure) would necessitate reducing the already short length of the panel, and the alternative ASHE-based measure was not used due to its worrying lack of correlation with the more accurate APS-based measure.

The results of the non-spatial models are presented below in Table 8. Each of the three models presented use fixed-effects and have standard errors that are robust to the fact the districts in the sample are observed multiple times over the years.

Table 8: Results of Non-spatial Models

	Total	Short-term	Long-term
Proportion of Working Age Population Aged 50 or Above	-0.000 (0.00)	-0.000 (0.00)	0.000 (0.00)
Proportion of Working Age Population Aged 16 to 24	-0.003 (0.00)	-0.001 (0.00)	-0.002 (0.00)
Working Age Participation Rate	0.004 (0.00)	0.002 (0.00)	0.002 (0.00)
Proportion of Working Age Population with a Degree or Above	-0.019*** (0.00)	-0.013*** (0.00)	-0.005*** (0.00)
Share of Employment in Largest Three Sectors	-0.000 (0.00)	-0.000 (0.00)	0.000 (0.00)
Natural Log of Median Gross Weekly Earnings	0.464** (0.17)	0.164 (0.11)	0.298*** (0.08)
House prices are in the middle third	-0.130* (0.05)	-0.099** (0.04)	-0.029 (0.02)
House prices are in the highest third	-0.124 (0.07)	-0.109* (0.05)	-0.012 (0.03)

Year is 2004	-1.175***	-0.611***	-0.563***
	(0.05)	(0.03)	(0.03)
Year is 2005	-1.165***	-0.578***	-0.586***
	(0.05)	(0.03)	(0.03)
Year is 2006	-0.987***	-0.448***	-0.540***
	(0.04)	(0.02)	(0.02)
Year is 2007	-1.188***	-0.624***	-0.563***
	(0.04)	(0.02)	(0.02)
Year is 2008	-1.119***	-0.464***	-0.653***
	(0.04)	(0.02)	(0.02)
Year is 2009	0.339***	0.896***	-0.555***
	(0.03)	(0.02)	(0.02)
Year is 2010	0.191***	0.525***	-0.335***
	(0.02)	(0.01)	(0.02)
Year is 2011	0.258***	0.629***	-0.371***
	(0.01)	(0.02)	(0.02)
Year is 2012	0.388***	0.418***	-0.029***
	(0.01)	(0.01)	(0.01)
Constant	0.572	1.525*	-0.933
	(1.09)	(0.72)	(0.49)
Observations	3790	3790	3790
R^2	0.80	0.84	0.66
AIC	2479	-88	-3070
ll	-1222.727	60.976	1552.119

Robust standard errors in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

From these results it is clear that national factors, incorporated through the inclusion of year-fixed effects, are very important for district unemployment. Amongst the other variables, human capital, house prices and wages appear to be important while the other factors included do not appear to be significantly associated with district unemployment rates.

These models do not allow for any spatial dependence so it is important to analyse whether this presents any problems. One issue could be that the residuals are spatially autocorrelated meaning that the assumption that they are independent of each other will not hold. This can be examined by calculating Moran's I values for the residuals as I have done in Table 9 which shows that the residuals from each of the three models are significantly positively spatially autocorrelated.

Table 9: Moran's I for Residuals of Non-spatial Models

	Moran's I	P value
Total Unemployment Residuals	0.42	0.00
Short-term Residuals	0.39	0.00
Long-term Residuals	0.43	0.00

The presence of this spatial autocorrelation in the residuals lends support for the estimation of spatial models, alongside the reasons already established such as my desire to identify spatial spillovers that affect unemployment. As mentioned my estimation of spatial models began with the SDM specification which was then subject to tests to provide support for its appropriateness. I also carried out a Hausman test to ensure fixed effects were statistically required as well as theoretically supported. The final testing I performed aimed to determine which spatial weight matrix provided the best fit and was done using comparisons of AIC values.

The first tests therefore focussed on the appropriateness of the SDM. The results of the tests[13] suggested that the spatially lagged dependent variable is significant at the 1% level as are the spatially lagged independent variables when considered jointly. This means the SDM was preferred to the SAR. The test comparing the SDM to SEM also supported the SDM as the suggestion that $\theta = -\rho\beta$ (which would reduce the SDM to the SEM, (Elhorst (2014)), was rejected at the 1% level.

The next element of the model that I tested was whether the decision to used the fixed effects variant of the SDM model was statistically necessary as well as being preferred on theoretical grounds. I achieved this using a Hausman test, the results of which[14] suggested that using a random effects model would yield inconsistent estimates, supporting my use of the fixed effects variant.

The final element I tested was the selection of the spatial weight matrix which best approximates the true data generating process. This was done using a comparison of the Akaike Information Criteria relating to the models. These AIC values can be seen in Table 10 and suggest that for all unemployment types, using the matrix with a 50km distance threshold results in a model which best approximates the true data generating process as this specification yields the lowest AIC.

[13] Available upon request
[14] Available upon request

Table 10: AIC Values Using Different Spatial Weight Matrices

Spatial Weight Matrix	Total Unemployment	Short-term Unemployment	Long-term Unemployment
25km	1308	-1140	-3915
50km	**914**	**-1593**	**-4232**
100km	1152	-1398	-4052
200km	1275	-1281	-3961
300km	1346	-1215	-3997
Contiguity	1248	-1234	-4114
Non-spatial model	2479	-88	-3070

Having established that the best model specification is a fixed effects SDM using a spatial weight matrix with a 50km distance threshold, it is now possible to present some results from this model specification. These results are outlined in the following tables which all refer to the same three models (one each for total, short-term and long-term unemployment). I have used separate tables for the different estimate types in order to ease presentation and make the results easier to digest.

The first set of estimates are coefficients on the independent variables (referred to in Table 5 as β). There are only two of the presented variables which are significant at conventional levels, the human capital and wage measures, although the majority of the year dummies are also significant but have not been shown. This is similar to the non-spatial models already presented.

As already noted earlier in the paper, the estimates from spatial regression models are best interpreted through the direct effect, indirect effect and total effect summary measures and therefore more interpretation can be found in the passages which deal with these measures.

Table 11: Independent Variable Coefficient Estimates

	Total	Short-term	Long-term
Independent Variables			
Proportion of Working Age Population Aged 50 or Above	0.000	0.000	-0.000
	(0.00)	(0.00)	(0.00)
Proportion of Working Age Population Aged 16 to 24	-0.002	-0.001	-0.001
	(0.00)	(0.00)	(0.00)
Working Age Participation Rate	0.002	0.001	0.001
	(0.00)	(0.00)	(0.00)
Proportion of Working Age with a Degree or Above	-0.006**	-0.006***	-0.001
	(0.00)	(0.00)	(0.00)

Share of Employment in Largest Three Sectors	0.002	0.001	0.001
	(0.00)	(0.00)	(0.00)
Natural Log of Median Gross Weekly Earnings	0.354**	0.135	0.215***
	(0.12)	(0.08)	(0.06)
House prices are in the middle third	-0.034	-0.033	-0.002
	(0.04)	(0.03)	(0.02)
House prices are in the highest third	0.019	-0.009	0.028
	(0.06)	(0.04)	(0.03)
Year Dummies	Included	Included	Included

Robust standard errors in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 12 includes the estimates of the coefficients on the spatially lagged independent variables (referred to in Table 5 as θ) and the spatially lagged dependent variable (referred to in Table 5 as ρ). The human capital measure is the only one of the independent variables that is significant at conventional levels (with the exception of one of the house price measure in one of the models) but the spatially lagged dependent variable is also significant with a magnitude just below 0.7 in each of the models.

Table 12: Spatially Lagged Independent and Dependent Variable Coefficient Estimates

Spatially Lagged Independent Variables	Total	Short-term	Long-term
Proportion of Working Age Population Aged 50 or Above	0.017	0.009	0.009
	(0.01)	(0.01)	(0.01)
Proportion of Working Age Population Aged 16 to 24	0.012	0.008	0.004
	(0.01)	(0.01)	(0.01)
Working Age Participation Rate	-0.001	-0.002	0.002
	(0.01)	(0.01)	(0.00)
Proportion of Working Age with a Degree or Above	-0.024***	-0.012*	-0.013***
	(0.01)	(0.00)	(0.00)
Share of Employment in Largest Three Sectors	-0.002	0.001	-0.003
	(0.01)	(0.00)	(0.00)
Natural Log of Median Gross Weekly Earnings	0.335	0.210	0.180
	(0.42)	(0.28)	(0.17)
House prices are in the middle third	-0.170	-0.106	-0.075*
	(0.09)	(0.07)	(0.04)

House prices are in the highest third	-0.172	-0.083	-0.095
	(0.12)	(0.09)	(0.05)
Spatially Lagged Dependent Variable			
Rho	0.693***	0.688***	0.663***
	(0.03)	(0.03)	(0.04)

Robust standard errors in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 13 includes the first of the summary measures which is the direct effect. This describes the effect of a unit change in an explanatory variable for one region on the dependent variable of that region, averaged over all the regions (LeSage and Pace (2009)). Two of the direct effects are significant, that of human capital and that of wages.

The magnitude of the effect of human capital suggests that a one percentage point increase in the proportion of working age residents with a degree or above in a district decreases the total unemployment rate in that district by 0.009 percentage points on average, other things being equal. The effect is also significant for short-term unemployment but not for long-term unemployment. The magnitudes of the estimated direct effects of human capital increases are larger (in absolute terms) than the coefficients shown in Table 11. This suggests that feedback effects included in the model (where the impact passes through neighbouring districts and back to the district in question) tend to reinforce the unemployment reducing effect of human capital improvements.

With regards to wages, a 1% increase in gross weekly earnings in a district increases the unemployment rate in that district by 0.004 percentage points on average, other things being equal. This effect is also larger in magnitude than the coefficients in Table 11 indicating that again there is reinforcing feedback. Another interesting point is that although the coefficient on wages in the short term-model is insignificant, the direct effect of wages is significant at the 5% level, highlighting the importance of looking at the estimated effects when dealing with a spatial model.

Table 13: Direct Effect Estimates

Direct Effects	Total	Short-term	Long-term
Proportion of Working Age Population Aged 50 or Above	0.002	0.001	0.001
	(0.00)	(0.00)	(0.00)
Proportion of Working Age Population Aged 16 to 24	-0.000	0.000	-0.001
	(0.00)	(0.00)	(0.00)
Working Age Participation Rate	0.002	0.000	0.001
	(0.00)	(0.00)	(0.00)
Proportion of Working Age with a Degree or Above	-0.009***	-0.007***	-0.002
	(0.00)	(0.00)	(0.00)

Share of Employment in Largest Three Sectors	0.002	0.001	0.001
	(0.00)	(0.00)	(0.00)
Natural Log of Median Gross Weekly Earnings	0.445***	0.185*	0.261***
	(0.13)	(0.08)	(0.06)
House prices are in the middle third	-0.056	-0.047	-0.010
	(0.05)	(0.03)	(0.02)
House prices are in the highest third	-0.002	-0.021	0.019
	(0.07)	(0.05)	(0.03)

Robust standard errors in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 14 presents the estimates of the indirect effects. These give the effect of a unit change in an explanatory variable in all regions other than one, on the dependent variable in that one region, averaged over all regions (LeSage and Pace (2009). There are two indirect effects that appear to be significant, that of human capital and that of house prices.

The estimates of the former effect suggest that a one percentage point increase in the proportion of working age residents with a degree or above in all districts apart from one decreases the total unemployment rate in that one district by 0.089 percentage points on average, other things being equal. This indirect effect is also significant for both short- and long-term unemployment and is substantially larger (for all types) than the direct effect.

Only one of the house price dummies is significant, and it is only significant for two of the three models. The magnitude of the coefficient in the total unemployment model suggests that on average, if house prices in all districts other than one increased from levels seen in the bottom third of the distribution to levels seen in the middle of the distribution this would be associated with an unemployment rate 0.612 percentage points lower in that one remaining district. One note of caution regarding the indirect effect of house prices comes from the fact that the indirect effects of both house price dummies are jointly insignificant.

Both of the significant indirect effects are larger in magnitude than the spatially lagged coefficient estimates (the θ values shown in Table 12). This again points to the existence of reinforcing feedback effects.

Table 14: Indirect Effect Estimates

Indirect Effects	Total	Short-term	Long-term
Proportion of Working Age Population Aged 50 or Above	0.056	0.028	0.025
	(0.03)	(0.02)	(0.02)
Proportion of Working Age Population Aged 16 to 24	0.040	0.028	0.009
	(0.04)	(0.02)	(0.01)
Working Age Participation Rate	0.003	-0.005	0.008
	(0.03)	(0.02)	(0.01)
Proportion of Working Age with a Degree or Above	-0.089***	-0.049***	-0.037***
	(0.02)	(0.01)	(0.01)
Share of Employment in Largest Three Sectors	-0.004	0.003	-0.005
	(0.02)	(0.01)	(0.01)
Natural Log of Median Gross Weekly Earnings	1.741	0.856	0.899
	(1.41)	(0.93)	(0.49)
House prices are in the middle third	-0.612*	-0.392	-0.216*
	(0.30)	(0.22)	(0.11)
House prices are in the highest third	-0.523	-0.280	-0.216
	(0.41)	(0.29)	(0.16)

Robust standard errors in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 15 provides the estimates of the total effects. These describe the effect of a unit change in an explanatory variable for all regions on the dependent variable in one region, averaged over all the regions (LeSage and Pace (2009)). Only three of the total effects are significant; human capital, wages and one of the house price variables (though as with the indirect effects, the total effects of the house price measures are jointly insignificant at the 5% level).

The total effects of wages and one of the house price variables are only significant for one and two of the models respectively. The total effect of the human capital measure, however, is significant for all three unemployment types and suggests that a one percentage point increase in the proportion of working age residents with a degree or above in all districts decreases the total unemployment rate in a given district by 0.098 percentage points on average, other things being equal.

Table 15: Total Effect Estimates and Model Details

Total Effects	Total	Short-term	Long-term
Proportion of Working Age Population Aged 50 or Above	0.058	0.029	0.026
	(0.03)	(0.02)	(0.02)

Proportion of Working Age Population Aged 16 to 24	0.040	0.029	0.007
	(0.04)	(0.03)	(0.01)
Working Age Participation Rate	0.005	-0.004	0.009
	(0.03)	(0.02)	(0.01)
Proportion of Working Age with a Degree or Above	-0.098***	-0.056***	-0.039***
	(0.02)	(0.02)	(0.01)
Share of Employment in Largest Three Sectors	-0.001	0.005	-0.004
	(0.02)	(0.01)	(0.01)
Natural Log of Median Gross Weekly Earnings	2.187	1.041	1.161*
	(1.45)	(0.96)	(0.50)
House prices are in the middle third	-0.669*	-0.438	-0.226*
	(0.32)	(0.24)	(0.11)
House prices are in the highest third	-0.524	-0.301	-0.197
	(0.45)	(0.31)	(0.17)
Observations	3790	3790	3790
AIC	914	-1593	-4232
ll	-379	874	2194

Robust standard errors in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In terms of how these findings compare to the non-spatial models I presented in Table 8, there is some similarity. For example, the non-spatial models also found significant effects from human capital and wages. In addition, the magnitudes of the estimated direct effects of wage increases in the spatial models are similar to the coefficient estimates in the non-spatial models. The estimated direct effects of human capital increases, however, are a little smaller than the relevant coefficients in the non-spatial model.

The indirect effects of the spatial model have no point of comparison in the non-spatial model and even though the direct effects can be compared with the coefficient estimates from the non-spatial model, the direct effect estimates include feedback. The ability to capture these spatial phenomena is one of the key advantages of using spatial models.

The models above have presented findings relating to total, short-term and long-term unemployment in order to facilitate comparisons between the latter two categories in particular. Based on the results obtained, however, there are not too many differences to be pointed out as the same spatial weight matrix produces the best model fit for both unemployment types, and for both a fixed effects SDM appears to be the most appropriate of the specifications considered.

The two minor differences present are that the direct effect of human capital increases only appears to be significant for short-term unemployment while the total effect of wage increases is only significant for long-term unemployment. The differences in the significance of one of the house price variables should not be given too much attention given the measurement problems in this variable (explained in section 3.2.6) and the fact that the two variables taken together are jointly insignificant.

In order to have a little more confidence in the findings presented in this chapter, I carried out a robustness check by comparing the existing results to those obtained when using the contiguity matrix. These exact estimation results can be found in Table 18 but overall these results should provide additional confidence in those from the main specification relating to both wages and human capital. This is because the results relating to these variables are very similar. This is not the case for the house price variables which were insignificant in all cases when the contiguity matrix was used. In the next section I discuss these seemingly robust results, aiming to explain them and use them to answer the remaining research questions. In addition the discussion section outlines my views regarding the main limitations of the study and offers my recommendations for policy and for future research.

6. Discussion

In the previous section I presented the main results from the exploratory and econometric analysis carried out. My exploratory analysis highlighted that unemployment in Great Britain is positively spatially autocorrelated and that there are four large clusters that particularly contribute to this spatial autocorrelation. My econometric analysis explored this in more detail by estimating SDMs for total, short-term and long-term unemployment, finding significant spatial spillovers affecting all three.

In this section I use the results obtained to answer my main research questions and constituent sub-questions before explaining the limitations of this study and what they mean for the answers given. The discussion section also includes my recommendations for policy and for future research.

6.1. Answers to the Remaining Research Questions

This study had one main research question and seven sub-questions. I answered the first four of these questions in the introduction and theoretical framework and built upon these initial answers to generate the evidence required to answer the remaining questions by carrying out exploratory and econometric analysis.

Of the remaining questions, the fifth asked whether there is spatial dependence in unemployment in Great Britain. The values of Moran's I produced in the exploratory analysis clearly demonstrated significant positive spatial autocorrelation and the spatial model of total unemployment estimated as part of the econometric analysis was able to capture significant spatial dependencies between the unemployment rates of the various districts. This suggests that I can answer the fifth research sub-question in the affirmative.

The same is also possible for the sixth and seventh research questions which asked, respectively, whether there is spatial dependence in short-term and long-term unemployment in Great Britain. As with total unemployment, both short- and long-term unemployment were positively spatially autocorrelated and were characterised by significant dependencies between districts. The answers to sub-questions five six and seven conform with the hypotheses I outlined at the end of section 2.7 and so these hypotheses should not be rejected.

This leads on to my main research question which asked whether there are differences in spatial dependence between short- and long-term unemployment. Across many of the comparisons there was little to separate short- and long-term unemployment as, for example, both had significant positive spatial autocorrelation of similar magnitudes and were best modelled using a SDM with a spatial weight matrix with a 50km distance threshold. Both models also demonstrated that there were significant indirect effects from human capital improvements and the existence of feedback effects as a result of spatial dependency.

Therefore, based on the findings in this study, there does not appear to be a difference in the spatial dependencies affecting short- and long-term unemployment so my hypothesis regarding the main research question should be rejected. In order to understand this answer and the answers to the fifth, sixth and seventh research questions, it is necessary to return to the arguments I set out in the theoretical framework.

In the theoretical framework section, I explained that there may be spatial dependence in unemployment as job seekers search for and gain employment in multiple districts, not just their district of residence. This means that unemployment in a given district depends on factors in multiple districts, not just that district itself.

Consequently, the finding that there is spatial dependence in short-term and long-term unemployment but that this spatial dependence does not differ between the two unemployment types suggests that both the short- and long-term unemployed do search for jobs in multiple regions under the influence of various factors, but that this study provides no evidence of this search behaviour varying between the groups.

The key factors found to significantly affect unemployment in the models included human capital and wages. These effects can also be better understood using the theoretical framework, for example, increases in the proportion of working age residents with a degree in a district were found to be associated with lower unemployment rates in that district. One potential explanation for this could be that individuals with more human capital carry out more effective job searches and are more likely to secure a position. This means they would have higher perceived chances of success and that districts with greater proportions of individuals with high levels of human capital are expected to have lower unemployment rates.

Similarly, it was found that higher proportions of graduates in neighbouring districts are negatively associated with the unemployment rate in a given district. An explanation for this could be that the productivity benefits of a concentration of skilled workers stimulate job growth. This would increase perceived chances of success for residents in nearby districts who can potentially access these jobs, meaning the districts in which they lived had lower unemployment. Other explanations for the beneficial indirect effect include the argument that high levels of human capital may signal that an area includes a large proportion of firms in knowledge intensive industries which could be positive for the attractiveness of jobs in this area, something which also encourages search activity and is therefore associated with lower unemployment.

The ability of my study to illuminate such spatial spillovers is one of the major contributions it has made as although past studies such as Patacchini and Zenou (2007) have examined the existence of spatial dependence in unemployment, they have placed much less focus on identifying these spatial spillovers. Another advantage of the work I have carried out is the comparison made between short- and long-term unemployment. Despite the fact there is no evidence of differences, this in itself is interesting given the theoretical reasons why there may have been differences.

My study is, of course, not without limitation and the next section aims to illuminate some of what I think are the key limitations. In doing this, the next section should help readers scrutinise my answers to the research questions and the recommendations which I have suggested based on the findings made.

6.2. Limitations of the Study

In order to understand how much faith should be placed in the answers to the research questions presented previously, it is necessary to understand some of the limitations of the study. The key limitations can be grouped into those related to the theoretical framework, those related to the data and those related to the methodology used.

The theoretical framework, while successful in providing a theoretical explanation for why there may be spatial dependence in unemployment, has the weakness of perhaps not being sufficiently explicit about the exact functional form of dependencies between different districts. Although I made theoretical arguments that search behaviour, which affects job matches and therefore unemployment, is multi-regional, I did not explicitly outline the functional form of the matching functions resulting from this multi-regional search activity. Instead I argued that job-matches in non-home regions were a function of distance to this region, attractiveness of working in the region and the probabilities of being successful from search activity in the region. This argument provided some justification for the functional form used, which included a distance-based spatial weight matrix and a SDM model with spatially lagged dependent and independent variables, but this justification was not as strong as if the functional form of the assumed matching functions had been stated.

This limitation was necessary as my theoretical framework could not have covered everything that was potentially of relevance, but has the implication that there is more uncertainty of whether the SDM was the best functional form than there might have been otherwise. The specification tests I carried out as part of the econometric analysis do provide confidence that the SDM is more appropriate than the SAR or the SEM, but of course a theoretical framework that explicitly considered the exact nature of the matching function could have suggested an alternative model, not within this set of three.

Another limitation of my theoretical framework is that there was little in the way of past academic research to support some of the theoretical indirect effects of the variables mentioned in section 2.6. This is because, as mentioned, few studies have placed as much focus on identifying spatial spillovers in unemployment at this study does so there were few for me to draw from.

There are a number of limitations that relate to the data I used. Firstly, the study was affected by the absence of some variables which could theoretically be influential such as housing tenure or environmental amenities. Secondly, the study had to use some variables which were not as robustly measured as would be desired. An example of this is the house price measure which was collected using slightly different procedures in Scotland compared to the rest of Great Britain. Thirdly, the study was constrained in the choice of the time periods and geographical units used as a result of data unavailability.

These limitations should not fundamentally undermine my findings but do suggest a relatively cautious approach, focussing on sign and significance, should be taken when using these findings as an input to decisions over policy or further research. There are two major reasons not to be excessively concerned by these data limitations which are my use of fixed-effects, which should control for time-invariant unobservable characteristics, and my use of the SDM, which LeSage and Pace (2009) shows (in section 3.3) reduces the bias resulting from spatially dependent omitted variables relative to OLS estimates.

Methodological limitations of the paper include the fact that the spatial weight matrix must be specified rather than estimated. This is a feature of spatial econometrics which is often criticised, for example see Elhorst (2010), but appears to be a necessary evil. This is because the 379 by 379 spatial weight matrices I used in this study included almost 150,000 elements (though of course many of these are zeroes) so would have been impossible to estimate.

The implication of this limitation is that the obtained results are sensitive to the choice of spatial weight matrix. This of course could lead to a 'data mining' type approach where the spatial weight matrix selected is the one that gives the 'desired' results and while the same can occur in non-spatial models (for example through adding, removing or transforming variables) it represents a particular limitation here too which necessitates caution when reading the results. I have attempted to minimise concerns of this character through selecting the matrix to be used objectively (using the AIC values) and by checking the results for robustness using an alternative spatial weight matrix. These attempts should convince readers that my findings obtained are of interest.

Related to this limitation regarding the spatial weight matrix are criticisms of the spatial econometrics literature as a whole such as Gibbons & Overman (2012), for example, which argues that studies using spatial econometrics are of limited use as they may face problems of identification (in part because the true form of the spatial weight matrix is not known). The implication of this criticism would be that less confidence can be placed in my results, however, there is a major reason why this criticism should not render this study unimportant. This is that I used panel models with fixed-effects, a strategy actually advocated in Gibbons & Overman (2012), for dealing with the afformentioned problems.

These limitations, which lessen but should not destroy, confidence in my findings were held in mind in the following two sections. In these sections I used my findings, alongside knowledge of my paper's limitations, to generate well grounded recommendations for policy and for future research.

6.3. Recommendations for Policy

My findings point to a number of recommendations for policy which stem both from the presence of spatial dependence and from the types of spillovers that have been detected. The presence of spatial dependence in unemployment suggests that public authorities at the district level should coordinate and cooperate with each other on any policies aiming to reduce unemployment. This is because the analysis in this study has shown than unemployment rates in a region depend on factors relating to other regions.

The results of the comparisons of models using different weight matrices which suggested the 50km distance threshold best approximated the true data generating process can be informative for districts in identifying which other areas appear particularly important for their fortunes. If the neighbourhood-set districts are faced with when using this threshold is too burdensome, for example around 60 districts would have over 50 neighbours using this definition, the similarity of results when the contiguity matrix suggests that focussing on collaborating with bordering regions may be an appropriate alternative.

The nature of the spatial spillovers detected suggests the second policy recommendation. This recommendation is to focus on improving human capital by increasing the proportion of workers with qualifications at degree level or above. The results indicate this would be a worthwhile policy as both the direct effects and indirect effects of human capital improvements appear to be beneficial for unemployment. Such a policy would be particularly attractive if it accompanied the regional collaboration already suggested or if it was a national policy as this would mean the beneficial spatial spillovers, which appear particularly important in reducing long-term unemployment, were fully reflected in decisions regarding the size of any investment to be made.

If investment decisions were made by districts in isolation then there may be under investment relative to the socially optimal level as public authorities would only focus on direct effects for their district. This would be unfortunate, particularly in the three high unemployment clusters seen in Figure 7, given the magnitude of the estimated indirect effects and because of the adverse consequences of unemployment for individuals and districts.

The fact that the results of my study are able to point to reasonable policy recommendations such as those discussed above reinforce the usefulness and societal relevance of the analysis I have carried out. My study also had academic relevance and was able to generate recommendations for future research which I have discussed in the next section.

6.4. Recommendations for Future Research

The findings I have made in this study have helped generate knowledge in some areas where it was lacking, however, the findings also point to additional research gaps which should be addressed in the future. My first recommendation would be for a scholar fully embedded in the literature on job-search and matching functions to derive a model of regional unemployment based on a multi-regional matching function with a functional form reflecting the leading ideas in this field. This would be beneficial as it would address one of the main limitations of this study and provide additional evidence on the question of whether the SDM model is appropriate in this context.

Another piece of research that would be useful in the future would be to examine the direct and indirect effects of human capital improvements on unemployment in more detail. For example it would be interesting to see if there are any apparent benefits from human capital improvements at lower ends of the qualification spectrum such as increasing the proportion of the workforce that has qualifications at A-Level or equivalent[15] . This would be interesting as the school leaving age has recently been increased to the age at which these qualifications are taken (UK Government, 2014).

Estimating similar models to those used in the study, but with the use of a travel time spatial weight matrix would also be interesting. This is because such a matrix may better reflect how the costs of distance affect job-search and therefore matching and unemployment. Results from models using travel time may also be able to point to parts of Great Britain where investments in transport could have beneficial economic effect by allowing advantageous spatial spillovers to reach underperforming areas.

These suggested future pieces of research, if carried out, would build upon the insights generated by my study. These insights are summarised in the conclusion which is the final section of the report. In the conclusion I aim to repeat the answers to the research questions which have been presented previously so that the messages of my study are clear.

7. Conclusion

In this study I aimed to answer one main research question which asked whether there were differences in spatial dependence in short- and long-term unemployment in Great Britain. The results of my analysis, which have provided strong evidence that there is spatial dependence in unemployment in Great Britain but no evidence of differences in these dependencies between short- and long-term unemployment, suggest that the answer to this question is no.

[15] A-Levels are advanced secondary level qualifications, usually obtained at age 18

This is not what I had expected when I outlined my hypothesis regarding this question in section 2.7 but is still of interest given the theoretical justifications which helped me generate this hypothesis. This justification was that as job-search behaviour can explain spatial dependence in unemployment, differences in this behaviour between the short- and long-term unemployed may lead to different dependencies being detected. The fact that I have rejected my hypothesis does not necessarily mean this argument was wrong but it does mean that my study can't provide empirical support for this argument.

My study can, however, provide some empirical support for my hypotheses regarding sub-questions V, VI and VII which were that spatial dependence was present in total, short-term and long-term unemployment respectively. This is because, for all three unemployment categories, the exploratory analysis demonstrated positive spatial autocorrelation while the econometric analysis provided support for the existence of spatial spillovers, for example, the beneficial spillovers related to human capital.

My study also provided some useful evidence on how best to approximate the dependencies between unemployment rates in different regions. This evidence suggested that assuming that dependencies exist between regions up to 50km in distance from each other appears to give a better approximation of reality than when 25km, 100km, 200km or 300km are used as thresholds.

The ability of my study to provide this useful evidence relied on the answers to the first four research questions. Demonstrating in the introduction that the spatial pattern of unemployment in Great Britain involved clustering of regions with similar unemployment rates and wide disparities between some of these clusters, helped me justify my use of spatial analysis.

My choice of specific analytical techniques was guided by the theoretical framework which established firstly that regions can experience different unemployment rates as they can experience different labour market developments. Secondly that there may be spatial dependence in unemployment due to the multi-regional job-search behaviour of job-seekers and thirdly that these dependencies may differ between short- and long-term unemployment as job search-behaviour may vary between these two groups.

Overall, my study has provided answers to my main research question and all of the underlying sub-questions. There are three main points for readers to take away from this study. Firstly, there is strong evidence of spatial dependence in unemployment in Great Britain. Secondly, these dependencies appear similar for both short- and long-term unemployment, and thirdly, their existence and nature makes them an important consideration when carrying out unemployment reduction policies, particularly those focussed on improving human capital.

8. References

Ahn, N., De La Rica, S., Ugidos, & A. (1999). Willingness to move for work and unemployment duration in Spain. *Economica*, 66, pp. 335-357.

Anselin, L. (2005). *Exploring Spatial Data with GeoDa: A Workbook.* Urbana: Spatial Analysis Laboratory, Department of Geography, University of Illinois,.

Aragon, Y., Haughton, D., Haughton, J., Leconte, E., Malin, E., Ruiz-Gazen, A., & Thomas-Agnan, C. (2003). Explaining the pattern of regional unemployment: the case of the Midi-Pyrenees region. *Papers in Regional Science*, 82, pp. 155-174.

Badinger, H., & Url, T. (2002). Determinants of regional unemployment: some evidence from Austria. *Regional Studies*, 36:9, pp. 977-988.

Belotti, F., Hughes, G., & Mortari, A. (2013). XSMLE - A Command to Estimate Spatial Panel Models in Stata. *2013 German Stata Users Group Meeting.* Potsdam.

Belotti, F., Hughes, G., & Mortari, A. (2013). *XSMLE: Stata module for spatial panel data models estimation*. Retrieved from IDEAS: http://ideas.repec.org/c/boc/bocode/s457610.html

Blanchard, O. (2006). European unemployment: the evolution of facts and ideas. *Economic Policy*, January 2006, pp. 5-59.

Blanchard, O., & Katz, L. (1992). Regional Evolutions. *Brookings Papers on Economic Activity*, 1:1992.

Brown, S., & Sessions, J. (1997). A profile of UK Unemployment: Regional versus demographic Influences. *Regional Studies*, Vol 31.4, pp. 351-366.

Brunello, G., Lupi, C., & Ordine, P. (2001). Widening differences in Italian regional unemployment. *Labour Economics*, 8 (2001) pp. 103-129.

Chapman, K. (2005). From 'growth centre' to 'cluster': restructuring, regional development, and the Teesside chemical industry. *Environment and Planning*, Vol. 37, pp. 597-615.

Chasco, C. (2013). Econometric theory and estimation methods of autoregressive spatial models. *3rd Regional Science Association of the Americas (RSAMERICAS) and the 5th SOCHER conference.* Arica, Chile.

Corrado, L., & Fingleton, B. (2012). Where is the economics in spatial econometrics? *Journal of Regional Science*, Vol. 52, No. 2, pp. 210-239.

Cracolici, M., Cuffaro, M., & Nijkamp, P. (2007). Geographical distribution of unemployment: An analysis of provincial differences in Italy. *Growth and Change*, 38, 4, pp. 649-670.

Debelle, G., & Vickery, J. (1998). Labour market adjustment: Evidence on interstate labour mobility. *Reserve Bank of Australia, Economic Research Department, Research Discussion Papers*, 9801.

Détang-Dessendre, C., & Gaigné, C. (2009). Unemployment duration, city size and the tightness of the labor market. *Regional Science and Urban Economics*, 39, pp. 266-276.

Drukker, D., Peng, H., Prucha, I., & Raciborski, R. (2012). *SPPACK: Stata module for cross-section spatial-autoregressive models*. Retrieved from IDEAS: http://ideas.repec.org/c/boc/bocode/s457245.html

Elhorst, P. (2003). Specification and estimation of spatial panel mata models. *International Regional Science Review*, 26, 3, pp. 244-268.

Elhorst, P. (2003). The mystery of regional unemployment differentials: Theoretical and empirical explanations. *Journal of Economic Surveys*, Vol. 17 No5. pp. 709-748.

Elhorst, P. (2009). Spatial Panel Data Models. In *Book Chapter.* Retrieved from http://www.researchgate.net/publication/226957388_Spatial_Panel_Data_Models

Elhorst, P. (2010). Applied spatial econometrics: Raising the bar. *Spatial Economic Analysis*, 5:1, pp. 9-28.

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

Fahr, R., & Sunde, U. (2002). Employment status, endogenous regional mobility, and spatial dependencies in labour markets. *IZA Discussion Paper Series*, No. 521.

Fischer, M., & Wang, J. (2011). *Spatial Data Analysis: Models, Methods and Techniques.* Heidelberg Dordrecht London New York: Springer.

Foote, C. (2007). Space and time in macroeconomic panel data: Young workers and state-level unemployment revisited. *Federal Reserve Bank of Boston Working Papers*, No. 07-10.

Gibbons, S., & Overman, H. (2012). Mostly Pointless Spatial Econometrics. *Journal of Regional Science*, Vol. 52, No.2, pp. 172-191.

Gilmartin, M., & Korobilis, D. (2012). On regional unemployment: An empirical examination of the determinants of geographical differentials in the UK. *Scottish Journal of Political Economy*, Vol. 59, No. 2, pp.179-195.

Hughes, G. (2012). Implementing procedures for spatial panel econometrics in Stata. *Spanish Stata User Group Meeting.*

Hujer, R., Blien, U., Caliendo, M., & Zeiss, C. (2002). Macroeconometric evaluation of active labour market policies in Germany. *IZA Discussion Papers*, No. 616.

International Labour Organization. (2014, May 27). *International Labour Organization*. Retrieved from Global Employment Trends 2014: Where is the Unemployment Rate the Highest?:

http://www.ilo.org/global/research/global-reports/global-employment-trends/2014/WCMS_233936/lang--en/index.htm

Jimeno, J., & Bentolila, S. (1998). Regional unemployment persistence (Spain 1976-1994). *Labour Economics*, 25-51.

Klinger, S., & Rothe, T. (2012). The impact of labour market reforms and economic performance on the matching of the short-term and the long-term unemployed. *Scottish Journal of Political Economy*, Vol. 59, No. 1, pp. 90-114.

Krueger, A., & Mueller, A. (2010). Job search and unemployment insurance: New evidence from time use data. *Journal of Public Economics*, 94, pp. 298-307.

LeSage, J., & Pace, R. (2009). *Introduction to Spatial Econometrics.* Boca Raton: Chapman & Hall.

Longhi, S., & Nijkamp, P. (2007). Forecasting regional labor market developments under spatial autocorrelation. *International Regional Science Review*, 30, pp.100-119.

Lottmann, F. (2012). Explaining regional unemployment differences in Germany: a spatial panel data analysis. *SFB 649 Discussion Papers*, 2012-026.

Lundin, E. (2013). Strategic interaction vs. regulatory compliance among regulated utilities: The Swedish water sector. *IFN Working Paper*, No. 998.

Mankiw, G. (2008). *Brief Principles of Macroeconomics, 5th Edition.* Mason: South-Western Cengage Learning.

Manning, A., & Petrongolo, B. (2011). How local are labour markets?: Evidence from a spatial job search model. *CEP Discussion Papers*, No 1101, December 2011.

McCann, P. (2013). *Modern Urban and Regional Economics, 2nd Edition.* Oxford: Oxford University Press.

Mitchell, W., & Bill, A. (2005). A spatial econometric analysis of the irreversibility of long term unemployment in Australia. *Centre of Full Employment and Equity Working Papers*, No. 05-05.

Nickell, S. (1997). Unemployment and labor market rigidities: Europe versus North America. *The Journal of Economic Perspectives*, Vol. 11 No.3 (Summer, 1997) pp. 55-74.

Nistor, A. (2009). Assessing the Effectiveness of Human Capital Investments on the Regional Unemployment Rate in the United States: 1990 and 2000. *International Regional Science Review*, 31:1, pp. 65-91.

Office for National Statistics. (2013). *Unemployment and The Claimant Count.* Newport: ONS.

Overman, H., & Puga, D. (2002). Unemployment clusters across Europe's regions and countries. *Economic Policy*, 17, 34, pp. 115-148.

Patacchini, E., & Zenou, Y. (2005). Spatial mismatch, transport mode and search decisions in England. *Journal of Urban Economics*, 58, pp. 62-90.

Patacchini, E., & Zenou, Y. (2007). Spatial dependence in local unemployment rates. *Journal of Economic Geography*, 7 (2007), pp. 169-191.

Rae, A. (2012). Spatial patterns of labour market deprivation in Scotland: Concentration, isolation and persistence. *Local Economy*, 27 (5-6) 593-609.

Rogers, C. (1997). Job search and unemployment duration: Implications for the spatial mismatch hypothesis. *Journal of Urban Economics*, 42, pp. 109-132.

Smith, T., & Zenou, Y. (2003). Spatial mismatch, search effort and urban spatial structure. *Journal of Urban Economics*, 54, pp. 129-156.

StataCorp. (2001). *STATA Technical Bulletin: March 2001.* Texas: StataCorp.

Stoll, M. (1999). Spatial job search, spatial mismatch, and the employment and wages of racial and ethnic groups in Los Angeles. *Journal of Urban Economics*, 46, pp. 129-155.

Taylor, J., & Bradley, S. (1997). Unemployment in Europe: A comparative analysis of regional disparities in Germany, Italy and the UK. *Kyklos*, Vol 50 - 1997 - Fasc 2, pp. 221-245.

Theodore, N. (2007). New Labour at work: long-term unemployment and the geography of opportunity. *Cambridge Journal of Economics*, 31, pp. 927-939.

UK Government. (2014). *School Leaving Age*. Retrieved from Gov.UK: https://www.gov.uk/know-when-you-can-leave-school

Vega, S., & Elhorst, P. (2013). Modelling regional labour market dynamics in space and time. *Papers in Regional Science*, doi:10.1111/pirs.12018.

Wasmer, E., & Zenou, Y. (2006). Equilibrium search uneployment with explicit spatial frictions. *Labour Economics*, 13, pp. 143-165.

Webster, D. (2005). Long-term unemployment, the invention of hysteresis and the misdiagnosis of structural unemployment in the UK. *Cambridge Journal of Economics*, 29, pp. 975-995.

9. Appendix

9.1. Data Sources

9.2. Appendix Tables

Table 16: Summary Statistics for Independent Variables

	Mean	S.D.	Min	Max	Count
Proportion of Working Age Residents Claiming any Benefit	10.53	3.64	3.50	26.40	3790
Proportion of Working Age Population Aged 50 or Above	29.69	5.69	10.00	54.00	3790
Proportion of Working Age Population Aged 16 to 24	17.44	3.39	6.00	33.00	3790
Working Age Participation Rate	77.60	4.48	58.90	92.80	3790
Proportion of Working Age with a Degree or Above	20.77	9.45	2.40	91.40	3790
Proportion of Working Age with No Qualifications	11.87	4.73	1.30	32.30	3790
Share of Employment in Largest Three Sectors	65.79	5.11	49.00	84.00	3790
Gross Weekly Income	372.59	70.43	182.14	903.90	3790
Natural Log of Median Gross Weekly Earnings	5.91	0.17	5.20	6.81	3790
Median House Price	174865.83	71465.61	38375.00	942500.00	3790
Natural Log of Median House Price	12.00	0.36	10.56	13.76	3790
House prices are in the lowest third	0.33	0.47	0.00	1.00	3790
House prices are in the middle third	0.33	0.47	0.00	1.00	3790
House prices are in the highest third	0.34	0.47	0.00	1.00	3790
Percentage Change in Employment (ASHE-based)	1.52	9.72	-44.00	64.00	3790
Percentage Change in Employment (APS-based)	0.47	8.86	-41.00	47.00	3411

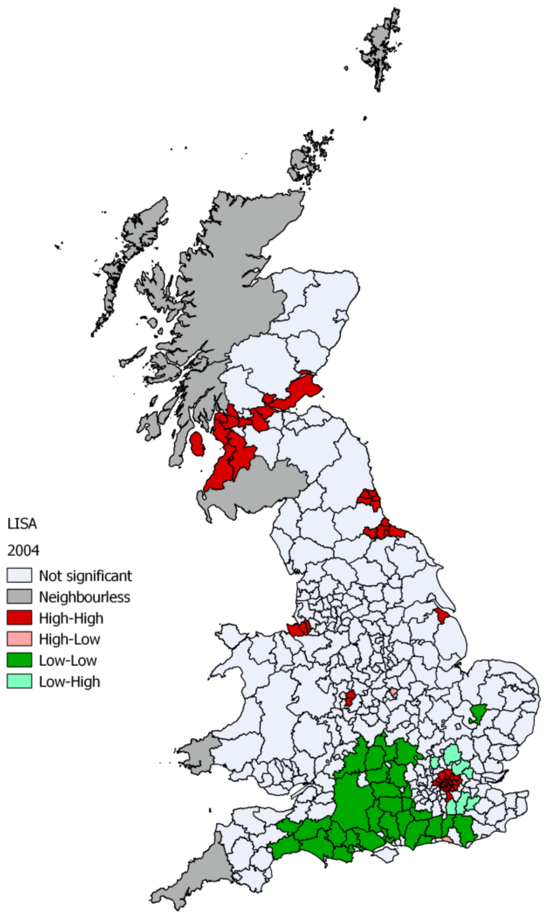
Source: ONS, various datasets as described in the main text. Some processing has been carried out by the author.

Table 17: Correlations

	Proportion of Working Age Residents Claiming any Benefit	Proportion of Working Age Population Aged 50 or Above	Proportion of Working Age Population Aged 16 to 24	Working Age Participation Rate	Proportion of Working Age with a Degree or Above	Proportion of Working Age with No Qualifications	Share of Employment in Largest Three Sectors
	1	**2**	**3**	**4**	**5**	**6**	**7**
1	1.00						
2	-0.12	1.00					
3	0.33	-0.37	1.00				
4	-0.65	0.14	-0.38	1.00			
5	-0.48	-0.29	-0.13	0.14	1.00		
6	0.65	-0.10	0.21	-0.52	-0.54	1.00	
7	0.34	-0.12	0.22	-0.31	0.07	0.09	1.00
8	-0.25	-0.44	0.01	0.04	0.60	-0.27	0.00
9	-0.70	-0.13	-0.27	0.28	0.71	-0.54	-0.15
10	0.67	-0.02	0.24	-0.36	-0.43	0.47	0.21
11	-0.14	0.15	-0.01	0.15	-0.12	-0.07	-0.10
12	-0.53	-0.13	-0.23	0.21	0.55	-0.40	-0.11
13	-0.02	-0.01	-0.03	0.02	-0.00	0.02	-0.01
14	-0.04	0.00	-0.02	0.11	0.02	-0.04	-0.01

	Natural Log of Median Gross Weekly Earnings	Natural Log of Median House Price	House prices are in the lowest third	House prices are in the middle third	House prices are in the highest third	Percentage Change in Employment (ASHE-based)	Percentage Change in Employment (APS-based)
	8	**9**	**10**	**11**	**12**	**13**	**14**
1							
2							
3							
4							
5							
6							
7							
8	1.00						
9	0.51	1.00					
10	-0.23	-0.74	1.00				
11	-0.16	-0.02	-0.50	1.00			
12	0.40	0.77	-0.50	-0.50	1.00		
13	-0.04	0.02	-0.03	-0.00	0.03	1.00	
14	0.02	0.04	-0.03	0.01	0.02	0.04	1.00

Figure 8: Total Unemployment Cluster Map (2004)



Source: Author's analysis of Office for National Statistics (ONS) and Ordinance Survey (OS) data

50km spatial weight matrix used. LISA stands for local index of spatial autocorrelation

Table 18: Model Results Using Contiguity-Based Spatial Weight Matrix

	Total	Short-term	Long-term
Independent Variables			
Proportion of Working Age Population Aged 50 or Above	-0.000 (0.00)	-0.000 (0.00)	-0.000 (0.00)
Proportion of Working Age Population Aged 16 to 24	-0.003 (0.00)	-0.002 (0.00)	-0.001 (0.00)
Working Age Participation Rate	0.003 (0.00)	0.001 (0.00)	0.002 (0.00)
Proportion of Working Age with a Degree or Above	-0.008*** (0.00)	-0.006*** (0.00)	-0.002* (0.00)
Share of Employment in Largest Three Sectors	0.001 (0.00)	0.000 (0.00)	0.000 (0.00)
Natural Log of Median Gross Weekly Earnings	0.360* (0.14)	0.121 (0.09)	0.234*** (0.06)
House prices are in the middle third	-0.038 (0.05)	-0.034 (0.04)	-0.003 (0.02)
House prices are in the highest third	-0.008 (0.07)	-0.024 (0.05)	0.017 (0.03)
Year Dummies	Included	Included	Included
Spatially Lagged Independent Variables			
Proportion of Working Age Population Aged 50 or Above	0.006 (0.01)	0.003 (0.00)	0.003 (0.00)
Proportion of Working Age Population Aged 16 to 24	0.003 (0.01)	0.000 (0.00)	0.003 (0.00)
Working Age Participation Rate	0.005 (0.00)	0.003 (0.00)	0.002 (0.00)
Proportion of Working Age with a Degree or Above	-0.025*** (0.01)	-0.016*** (0.00)	-0.010*** (0.00)
Share of Employment in Largest Three Sectors	-0.004 (0.00)	-0.002 (0.00)	-0.002 (0.00)

Natural Log of Median Gross Weekly Earnings	-0.223 (0.30)	-0.052 (0.20)	-0.142 (0.13)
House prices are in the middle third	-0.104 (0.16)	-0.044 (0.11)	-0.070 (0.05)
House prices are in the highest third	-0.150 (0.17)	-0.083 (0.12)	-0.080 (0.06)
Spatially Lagged Dependent Variable			
Rho	0.603*** (0.04)	0.584*** (0.03)	0.607*** (0.04)
Direct Effects			
Proportion of Working Age Population Aged 50 or Above	0.000 (0.00)	0.000 (0.00)	0.000 (0.00)
Proportion of Working Age Population Aged 16 to 24	-0.002 (0.00)	-0.001 (0.00)	-0.001 (0.00)
Working Age Participation Rate	0.004 (0.00)	0.002 (0.00)	0.002 (0.00)
Proportion of Working Age with a Degree or Above	-0.013*** (0.00)	-0.010*** (0.00)	-0.004** (0.00)
Share of Employment in Largest Three Sectors	0.001 (0.00)	0.000 (0.00)	0.000 (0.00)
Natural Log of Median Gross Weekly Earnings	0.385* (0.15)	0.141 (0.10)	0.247*** (0.07)
House prices are in the middle third	-0.058 (0.07)	-0.044 (0.05)	-0.015 (0.02)
House prices are in the highest third	-0.035 (0.09)	-0.041 (0.06)	0.004 (0.03)
Indirect Effects			
Proportion of Working Age Population Aged 50 or Above	0.013 (0.01)	0.006 (0.01)	0.007 (0.01)
Proportion of Working Age Population Aged 16 to 24	0.005 (0.02)	-0.000 (0.01)	0.005 (0.01)
Working Age Participation	0.015	0.007	0.008

Rate	(0.01)	(0.01)	(0.01)
Proportion of Working Age with a Degree or Above	-0.071*** (0.01)	-0.044*** (0.01)	-0.026*** (0.01)
Share of Employment in Largest Three Sectors	-0.008 (0.01)	-0.004 (0.01)	-0.003 (0.00)
Natural Log of Median Gross Weekly Earnings	-0.041 (0.70)	0.014 (0.45)	0.005 (0.30)
House prices are in the middle third	-0.295 (0.39)	-0.143 (0.27)	-0.170 (0.12)
House prices are in the highest third	-0.369 (0.42)	-0.227 (0.29)	-0.167 (0.15)
Total Effects			
Proportion of Working Age Population Aged 50 or Above	0.014 (0.01)	0.006 (0.01)	0.008 (0.01)
Proportion of Working Age Population Aged 16 to 24	0.003 (0.02)	-0.002 (0.01)	0.004 (0.01)
Working Age Participation Rate	0.019 (0.02)	0.009 (0.01)	0.010 (0.01)
Proportion of Working Age with a Degree or Above	-0.085*** (0.01)	-0.054*** (0.01)	-0.030*** (0.01)
Share of Employment in Largest Three Sectors	-0.008 (0.01)	-0.004 (0.01)	-0.003 (0.00)
Natural Log of Median Gross Weekly Earnings	0.344 (0.77)	0.155 (0.50)	0.252 (0.33)
House prices are in the middle third	-0.353 (0.44)	-0.187 (0.31)	-0.186 (0.14)
House prices are in the highest third	-0.403 (0.48)	-0.267 (0.34)	-0.162 (0.17)
Observations	3790	3790	3790
AIC	1248	-1234	-4114
ll	-546.156	695.140	2135.143