University of Tartu
Faculty of Social Sciences
School of Economics and Business Administration

SPATIAL INTERACTIONS OF REGIONAL LABOUR MARKETS IN EUROPE

Laura Helena Kivi

Tartu 2019

ISSN-L 1406-5967
ISSN 1736-8995
ISBN 978-9985-4-1139-1 (pdf)
The University of Tartu FEBA
https://majandus.ut.ee/en/research/workingpapers

Spatial interactions of regional labour markets in Europe

Laura Helena Kivi*

Abstract

This study investigates the spatial dependence of unemployment and employment rates in Europe relying on Eurostat NUTS2 level data for 306 European regions. Spatial dependence is explored using spatial error, spatial lag and a spatial autoregressive model with spatial autoregressive disturbances. The findings show that regional labour markets in Europe cluster in space – regions with high (low) (un)employment rate are surrounded by regions with high (low) (un)employment rate. The study provides evidence that significant spillovers across regional labour markets exist. The (un)employment rate in one region is directly affected by (un)employment rate changes in other regions, but also by unobserved shocks in other regions. It was found that the spatial effects are not determined by differences in the share of the population of youth, differences in industrial structure or difference in human capital. The results of the study confirm the importance of close coordination between regions while developing labour market and regional policy measures.

JEL Classification: C21, E24, R23

Keywords: regional labour markets, spatial econometrics, spatial dependence, clustering, Europe

*Laura Helena Kivi, PhD Student, School of Economics and Business Administration, University of Tartu, Estonia; researcher, Estonian Center for Applied Research (CentAR). E-mail: laura.kivi@centar.ee
This paper is based on my master thesis. I would like to thank my supervisor professor Tiiu Paas for her support and guidance throughout the whole writing process. I am grateful to Jaan Masso and Kadri Ukrainski for useful comments and thoughts on the earlier versions, to Amaresh K. Tiwari for his thoughtful comments and to Sven-Kristjan Bormann for his thorough feedback. I am solely responsible for all remaining errors and omissions.

1. INTRODUCTION

Regional disparities in labour market developments between different European countries have long been noted. While institutional differences between different countries might explain the disparities in unemployment and employment rates at country level, there still exists large differences in the given rates between regions within the same country. Based on the Eurostat database, regional unemployment rates have varied remarkably during the last decade. For instance, in 2015 in the case of Spain, from 13.8% to 34%, in Germany from 2.5% to 9.4% and in Italy from 3.8% to 22.9%. Beyer and Stemmer (2016) state that after the convergence in regional unemployment rates in Europe between 1996 and 2007, a polarization has followed in the period 2007–2013. Such large disparities in unemployment and employment result in differences in the income of individuals by region, leading to higher inequality between regions. Furthermore, as Taylor (1996) states, the reduction of regional differences in unemployment would lead to desired macroeconomic outcomes, such as higher national output and lower inflation.

Regional disparities have been investigated intensively in terms of unemployment rates, while employment rate disparities have received less attention. Many studies have tried to explain the regional variability of unemployment using differences in various factors, such as demographic factors, human capital, amenities, industrial composition, unemployment benefits (see Elhorst, 2003 for an overview). However, labour market participants are not restricted to work only in the region they reside in. While looking for employment opportunities, workers also consider neighbouring labour markets. Overman and Puga (2002) state that regional unemployment is related much more to neighbouring regions than to more distant regions within the same country. Positive spatial dependence in regional unemployment rates has been noted for different countries (Semerikova 2015, López-Bazo *et al.* 2002, Aragon *et al.* 2003, Filiztekin 2009, Cracolici *et al.* 2007). For instance, Badinger and Url (2002) report that spatial effects account for about one-fifth of the variation in the unemployment rate. Therefore, previous empirical evidence indicates that while investigating regional labour market differentials, it is essential to investigate the role of spatial dependence.

The aim of this study is to investigate spatial dependence at the regional level in Europe both in terms of unemployment and employment rates. To the best of the author's knowledge, no previous studies have been conducted where the spatial dependence of both unemployment and employment rates have been analysed with spatial models for the same dataset. It is important to investigate the spatial dependence of both the unemployment and employment rates to better understand the processes of regional labour markets. While both unemployment and employment are related to overall equilibrium in the labour market, their spatial dependence is likely to differ in terms of the strength of the relationship and possibly also their sign, as earlier studies have found mixed results in terms of the employment rate (see e.g. Pavlyuk 2011, Lewis *et al.* 2011, Mayor and López 2008). In the case where commuting and migration lead to positive spatial dependence in unemployment rates (see e.g. Molho 1995, Pattacchini and Zenou 2007), but competition for labour force among regions leads to negative dependence in employment rates, different labour market policy measures should be implemented to reduce the regional differences in employment and unemployment. The results of this study provide additional information for developing new labour market, social and regional policy measures in Europe taking into account possible spatial relations between the labour markets of countries and regions.

This study uses the European NUTS 2 regional data on unemployment and employment rates. The analysis is based on the results of spatial regression models. Specifically, the spatial lag model, spatial error model, spatial autoregressive model with spatial autoregressive disturbances and the spatial Durbin model are used to account for the spatial dependence in unemployment and employment rates. Demographic factors, industry structure, human capital and country dummies are added as explanatory variables.

The findings indicate that regional labour markets in Europe cluster in space; that is, regions with a high (low) (un)employment rate are surrounded by regions with a high (low) (un)employment rate. The results confirm the positive spatial dependence of unemployment and employment rates, even after controlling for regional characteristics. (Un)employment rate in one region is directly affected by (un)employment rate changes in other regions, but also by unobserved shocks in other regions. Significant spillovers exist across regional labour markets. Interestingly, spatial dependence between regional labour markets in Europe has been fairly stable during the last decade. No evidence was found that the spatial effects work through differences in demographics, such as the share of the population of youth, differences in the industrial structure or differences in human capital.

The paper is structured as follows. Section 2 presents the literature review. Section 3 explains the method and data. The empirical results are reported and discussed in Section 4. Robustness checks are presented in Section 5. Finally, Section 6 concludes.

2. LITERATURE REVIEW

In this section the focus is first on the theoretical framework for investigating labour market differentials. The equilibrium and disequilibrium views (Marston 1985) that explain the disparities of regional labour markets are examined. Factors that affect the adjustment of the regional labour market according to those views are analysed. While the factors affecting the unemployment rate are noted based on theoretical considerations, the results from empirical studies concerning the impact of those factors are also included.[1] In the second subsection, attention is turned specifically to the analyses of spatial dependence. Spatial dependence in labour markets has been mostly studied in the case of unemployment rates. Very few studies have been conducted with employment rates. Therefore, the first part of the second subsection deals with spatial dependence in unemployment rates and the background mechanism, while the second part focuses on the spatial dependence in the employment rate.

2.1. Theoretical framework of regional labour market disparities

There are many different factors causing unemployment rate differentials across regions. Marston (1985) states two possible explanations of the existence of disparities in unemployment rates: the equilibrium and disequilibrium view. Other studies (e.g. Aragon *et al.* 2003, Diaz 2006, Semerikova 2015) have followed his idea and added both disequilibrium and equilibrium based

[1] All of the empirical studies considered in this chapter (except Marston 1985) have accounted for spatial dependence between the regions by using different spatial econometric methods. This selection has been made on purpose, in order to be able to compare the results of this study to the previous analysis.

factors to their analysis to determine whether regional unemployment is more of a disequilibrium or equilibrium nature.

According to Marston (1985), in the equilibrium view each of the regions has its own stable long-run mean equilibrium unemployment rate. Although this underlying mean unemployment rate differs across regions, the distribution of rates is characterized by constant utility across regions. That means that a high unemployment rate in a region is compensated by some other factors (higher wages, amenities, lower overall cost of living, industry composition). In this view, external shocks affect unemployment rates only for a short period, allowing them to converge back to their underlying mean value (Semerikova 2015). Marston (1985) claims that if unemployment is more equilibrium by nature, then government attempts to reduce regional disparities are useless, as it is impossible to reduce the regional unemployment rate in the long term.

In the equilibrium view, most of the factors affecting the unemployment rate are variables that compensate for the high level of unemployment. Those variables are wages, amenities and industrial composition.

Traditionally, it is assumed that a rise in wages increases the unemployment rate as it decreases labour demand and increases labour supply. In the equilibrium view, the relationship is also predicted to be positive, as higher unemployment in the area is assumed to be compensated by higher wages in the area. The empirical results of Semerikova (2015), Aragon *et al.* (2003) and Marston (1985) support the equilibrium view of an average wage. On the contrary, Badinger and Url (2002) report a negative relationship, which is explained by the fact that opportunity costs to stay unemployed are higher in an area with a higher average wage.

Unemployment disparities may also originate from differences in amenities. As Aragon *et al.* (2003) state, according to the equilibrium view, areas with a more pleasant climate, active cultural life or better infrastructure are expected to exhibit higher unemployment. Lower housing costs are also sometimes seen as a compensation factor for high unemployment (Semerikova 2015). Population density is sometimes included as a measure of the quality of life. Areas with low population density can be seen as more favourable living environments as they tend to have stronger social networks (Badinger and Url 2002). Aragon *et al.* (2003), however, report a positive relationship for French data, but claim that densely populated urban regions can be considered more interesting and stimulating places to live in, in which case the finding fits the equilibrium view. Population density and share of urban areas can also affect the speed of adjustment of the labour market, an argument that will be considered under the disequilibrium view below.

When the composition of industries varies this is often seen as a factor of differentials in regional unemployment rates. Regions specialized in declining industries such as agriculture and manufacturing are assumed to exhibit higher unemployment rates than regions specialized in growing industries (Elhorst 2003). Often, shares of different industries in employment are used as controls. While some studies (e.g. Aragon *et al.* 2003) confirm the argument, most of the results are mixed or lacking significance (e.g. Semerikova 2015, López-Bazo *et al.* 2002, López-Hernández 2013, Filiztekin 2009, Niebuhr 2003, Diaz 2016).

In the disequilibrium view regional unemployment rates should become equal between the regions in the long run (Aragon *et al.* 2003). Workers from areas of high unemployment would migrate to other regions and firms would relocate to high unemployment areas seeking workforce, thereby levelling out the regional differences. However, the speed of the adjustment tends to be slowed down by restrictions on mobility on both sides: workers experience the cost of migration (e.g. housing costs) and firms are restricted by labour market rigidities (e.g. taxation, labour laws, welfare state arrangements, union agreements) (Marston 1985, Diaz 2016). Therefore, labour markets do not manage to reach this equal unemployment rate before a new shock (e.g. a factory closure) hits the labour market (Diaz 2016). Contrary to the equilibrium view, where long-run differentials could not be reduced by government policies, introducing more flexibility to labour markets and reducing migration costs could here help increase the speed of adjustment, hence reducing disparities in regional unemployment rates in the long term.

Under the disequilibrium view, the main variables affecting regional unemployment rates are those affecting the speed of adjustment. Those variables are age structure, average education level, employment growth, population density and the structure of the housing market.

Age structure of the population is thought be important in terms of adjustment. Young people are more likely to move to another region as their opportunity costs from moving are lower and they are less risk averse than older generations (Aragon *et al.* 2003). Filiztekin (2009) and Diaz (2016) confirm, based on regional data from Turkey and Colombia respectively, that the share of young people in the working age population is negatively related to unemployment. However, some studies also find that regions with a higher share of young people tend to have a more serious unemployment problem (López-Bazo *et al.* 2002, Mitchell and Bill 2004, Semerikova 2015). These findings might result from the barriers the younger generation has in terms of entering the labour market. The effect of the share of the older generation can be related to educational mismatch and constant changes in the industrial structure. Overall, the effect of age structure seems to be ambiguous.

Labour markets with more educated people tend to have lower unemployment rates for many reasons. First, the labour market for skilled workers tends to be geographically larger and their pay-off from moving is bigger, as they are potentially high-wage earners (Aragon *et al.* 2003). Highly skilled workers are also likely to be better informed and more efficient in finding jobs (Semerikova 2015). Lastly, highly educated are in greater demand in the labour market, and therefore they have greater opportunities to migrate (Elhorst 2003). Those theoretical considerations are in line with empirical results from Overman and Puga (2002), Diaz (2016), Marston (1985), López-Bazo *et al.* (2002) and López-Hernández (2013). Semerikova (2015) finds mixed results using German data as both people without a professional education and those with a university education have a positive effect on unemployment. Badinger and Url (2002) report no significant effect of skills structure on regional unemployment rates in Austria. Overall, based on theoretical considerations and empirical research, where a negative effect dominates, one would expect the higher share of high-skilled individuals to increase the speed of adjustment and lower the regional unemployment rates.

Employment growth reduces unemployment by definition, as it increases the labour force and might decrease the number of unemployed (a new worker might also come from non-participation or be a job migrant). This negative effect of employment growth is reported in most studies (e.g. Badinger and Url 2002, Diaz 2016, López-Bazo *et al.* 2002, Mitchell and Bill 2004, Niebuhr 2003, Semerikova 2015).

While population density and share of urban areas can be seen as one of the amenities in the equilibrium view, it is likely to affect the speed of adjustment and should therefore be considered also under the disequilibrium view. On one hand, job searching and matching in urban and more densely populated areas is faster and more efficient than in remote areas (Diaz 2016). On the other hand, urban areas attract job seekers from other regions and the accompanying supply effect might increase unemployment (Mitchell and Bill 2004). Semerikova's (2015) results for Germany support the former view, Niebuhr's (2003) results on European NUTS3 regions, the latter view.

The mobility of the workers, and therefore the speed of adjustment in the labour market is restricted by the magnitude of migration costs. Most important in this aspect is the structure of the housing market – housing prices and share of apartments owned/rented out. Badinger and Url (2002) find that regions with a higher share of public housing tend to have higher unemployment rates. People living in public housing experience a lock-in effect: they are afraid to give up their rental contract as the probability of finding other accommodation with subsidized cheaper prices, is low.

Overall, in the equilibrium view, labour market disparities between regions remain in the long run and high unemployment is compensated for by some other regional characteristics. In the disequilibrium view, disparities between regions diminish in the long run and might disappear eventually, depending on the speed of adjustment. The equilibrium view focuses on compensation factors, such as wages, amenities and industrial composition, while the disequilibrium view draws attention to factors affecting the speed of adjustment, such as the age structure and skill composition of the population, employment growth, population density and the structure of the housing market.

2.2. Spatial dependence in regional labour markets

While the abovementioned factors have an important role in explaining regional unemployment differentials, spatial dependence is found to be important as well. Badinger and Url (2002) report that spatial effects account for about one-fifth of the variation in the unemployment rate. Spatial dependence in regional unemployment rates has been shown to exist in Germany (Semerikova 2015), Japan (Kondo 2015), UK (Molho 1995, Pattacchini and Zenou 2007), Western Europe (Niebuhr 2003), Spain (López-Bazo *et al.* 2002), Turkey (Filiztekin 2009), Australia (Mitchell and Bill 2004), Colombia (Diaz 2016), France (Aragon *et al.* 2003) and Italy (Cracolici *et al.* 2007). In all of the cases spatial autocorrelation was positive, meaning that the neighbours of regions with high (low) unemployment rates also tend to have high (low) unemployment. Hence, regions tend to cluster in space in terms of their unemployment rates. It could be argued that detecting significant spatial autocorrelation in unemployment rates simply reflects the fact that neighbouring regions have similar local characteristics; for example, in terms of the skill composition of the population or industrial structure. However, almost all of the named studies (except Kondo 2015, and Pattacchini and Zenou 2007) also analysed spatial regression models, adding different controls to account for the various local characteristics, and still found significant spatial effects.

Although the studies mentioned above point to the existence of spatial dependence in unemployment rates, the underlining mechanism causing this dependence has not been identified in most of these studies. The main mechanisms that seem to cause spatial dependence are commuting and migration across neighbouring regions as people look for work both in the area they live in and in the areas they do not. An important contribution to the studies of this aspect has been made based on UK data by Molho (1995), who analyses the effects of supply and demand side shocks for regional unemployment taking spatial aspects into account. There are significant spillovers on adjustments to local demand shocks (in the form of employment growth) over a wider spatial field. At the time of the employment growth shock local unemployment is strongly affected but there are also small spillover effects on neighbouring areas that increase over time. The fact that the spillover effect is stronger after a time lag points to migration behaviour: when a higher labour demand is noted in the neighbouring labour market, workers need time to make arrangements (e.g. find appropriate housing, school for their children etc.) for the relocation. Molho (1995) also identifies highly localized effects that point to commuting. In line with those results is the study by Pattacchini and Zenou (2007), who focus on studying the commuting flows for UK Travel-To-Work-Areas. The authors find that spatial dependence is characterized by a low distance decay which points to the commuting behaviour of workers.

Few studies have investigated the spatial dependence in employment rates. Pavlyuk (2011) studies Latvian regional employment rates and finds the spatial lag to be negative. The somewhat surprising negative relationship seems to reflect the fact that there is competition among the regions for labour resources. It should be noted that the study uses geographically relatively small regions, which also might affect the results. Lewis *et al.* (2011) focus their analysis on spatial dependence in the manufacturing sector in the counties of South Carolina. Changes in manufacturing employment is found to have a positive relationship with employment changes in neighbouring counties. Here the sign of the relationship is likely to result from some positive cooperation effects among industries in different counties. Mayor and López (2008) use the employment data for NUTS 3 regions in Spain and, contrary to Lewis *et al.*, report the effects of the spatial dependence of employment change to be slightly negative. The results of studies of the spatial dependence of employment are therefore mixed in terms of the sign of the dependence. The mixed results might be explained by differences in data; for example, the differences in the size of the geographic units used in different studies or that they focus only on employment in one industrial sector (e.g. manufacturing sector). Alternatively, the results might indicate different forms of spatial interaction; for example, competition among regions for qualified workers resulting in a negative effect versus the agglomeration and cooperation of industries in different regions resulting in a positive effect.

In summary, previous studies have found positive spatial dependence in regional unemployment rates and somewhat mixed results on the sign of dependence in the case of employment change and rates. There is a clear gap in the literature in terms of employment rate spatial dependence. Furthermore, although there are rather numerous studies that analyse spatial dependence at the regional level for one country, the most recent study using spatial regression models to analyse regional unemployment spatial dependence for many neighbouring countries was Niebuhr (2003), using the 1986 and 2000 data on European NUTS 2 and NUTS 3 regions. Studying regional spillover not only within a country, but also between countries is important, while in light of the EU principle of the free movement of labour, workers in border regions are also likely to seek work opportunities in neighbouring regions across the national border.

This study focuses on the empirical analysis of spatial dependence in regional labour markets both for employment and unemployment rates and for regions in European countries. A spatial econometrics modelling approach is applied here in order to account for the inter-regional differences.

3. DATA AND METHOD

3.1. Data

The data used in the current study is provided by the Eurostat database. The regional unemployment rate is defined as a ration of the number of unemployed persons to the number of persons in the economically active population (i.e. sum of employed and unemployed). Unemployed persons comprise persons aged 15–74 who were: 1) without work during the reference week; 2) currently available for work; 3) actively seeking work or who had found a job to start within at most three months. The employed persons are those aged 15–64, who during the reference week did any work for pay, profit or family gain for at least one hour, or were not at work but had a job or business from which they were temporarily absent. The regional employment rate is defined as the number of employed persons in the population aged 15–64 (i.e. working age population). The indicators are based on the EU Labour Force Survey and in accordance with the International Labour Organization (ILO) definition of unemployment.[2]

The data on NUTS 2 level regions is used in the current study. The NUTS classification (Nomenclature of territorial units for statistics) is a system set up by Eurostat that establishes a hierarchy of three NUTS levels for each EU member state. The NUTS 2 level is defined as basic regions for the application of regional policies (NUTS overview). The NUTS 2013 classification is used in the current study. The study uses cross section data for 306 regions in Europe for the year 2015.[3]

Figure 1 presents the unemployment rates in European NUTS 2 regions. It can be seen that regions with similar unemployment rates are rather concentrated. Although most of the clustering seems to be within national borders (e.g. high unemployment in the south of Italy in comparison with central and northern Italy), some cross-border similarities can also be seen. For example, border regions in the south of Germany, Switzerland, Austria and the Czech Republic have similar low unemployment rates. Whether this is a sign of cross-border interaction between the regional labour markets or something that can be explained by regional similarities in terms of industrial structure, demographics and other regional variables, is something to be investigated by the following regression analysis.

Employment rates in NUTS 2 regions in Europe are displayed in Figure 2. Similar to unemployment, clustering can be seen inside national borders (e.g. Spain, France), but also across national borders. In accordance with the results on unemployment rates, regions in the south of Germany have similar employment rate values with their neighbours across the border in Switzerland and Austria. Again, the similarities in neighbouring regions seen in the raw data can be partially a sign of interaction across regions and partially accounted for by cross-country differences or regional characteristics (e.g. demographics, industrial structure).

[2] Description of the variables and source of exact datasets used are given in appendix 1.

[3] List of countries and number of regions included in the analysis is shown in appendix 2.

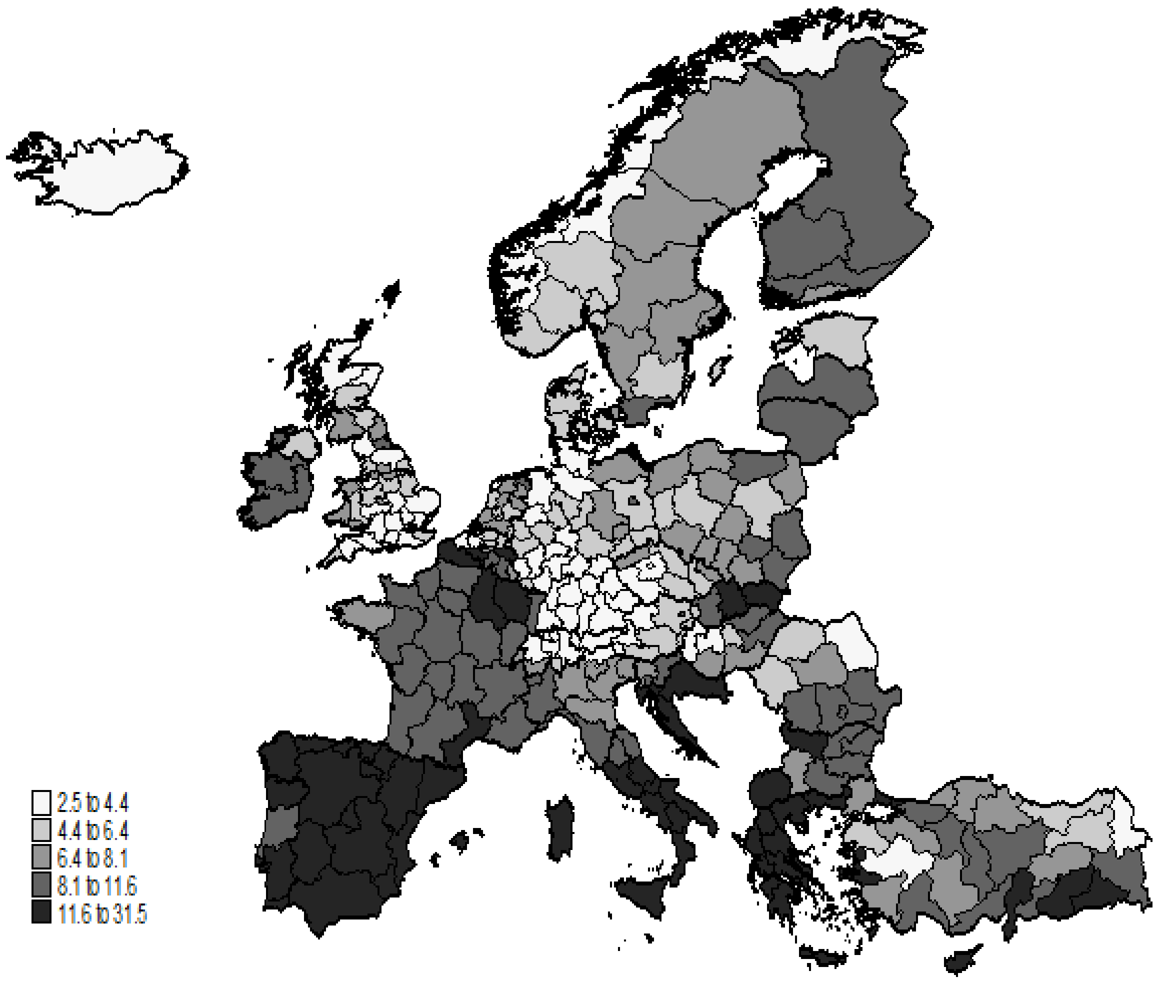

Figure 1. Regional unemployment rates (%) in NUTS 2 regions in Europe in 2015 (Eurostat database; compiled by the author)

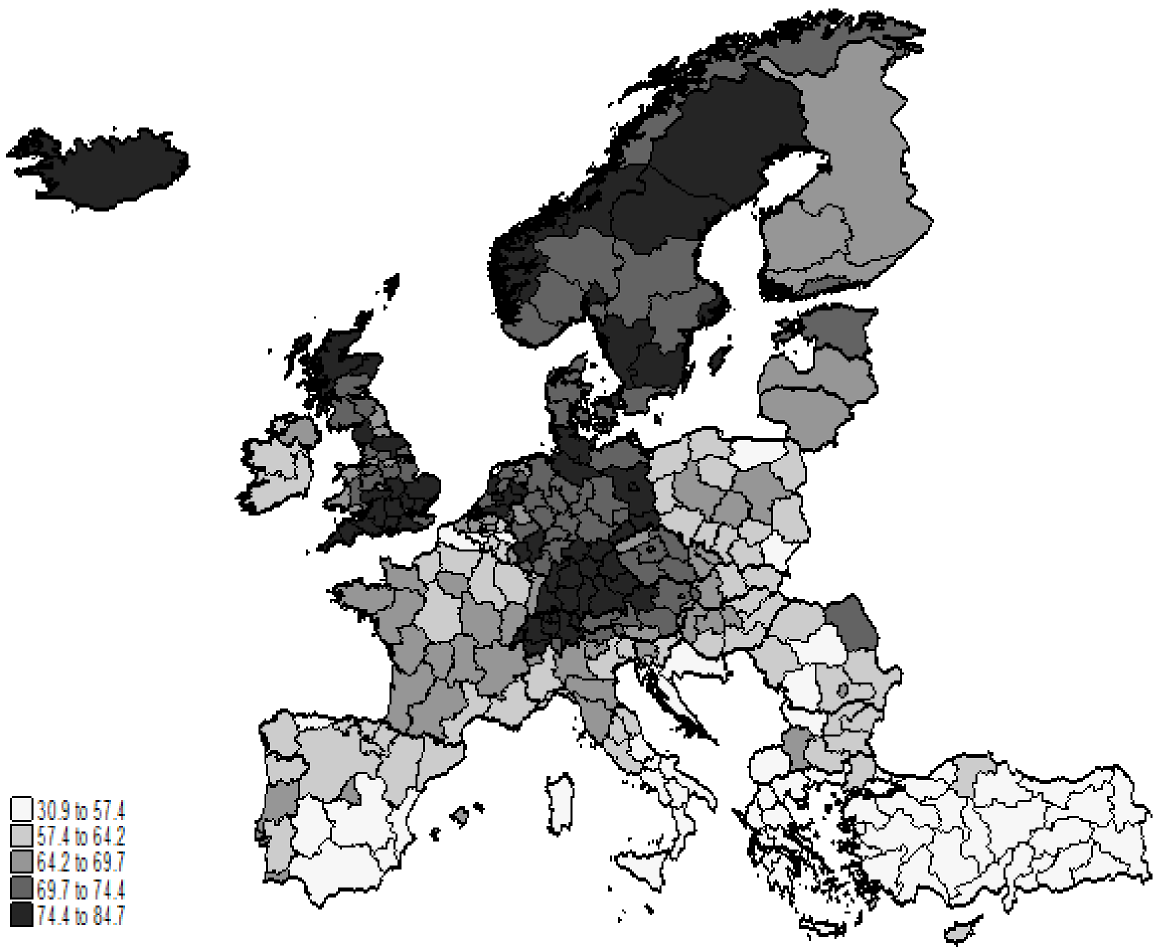

Figure 2. Regional employment rates (%) in NUTS 2 regions in Europe in 2015 (Eurostat database; compiled by the author)

According to the theory presented in the previous section, unemployment can be driven by both equilibrium and disequilibrium effects. To account for both effects, the explanatory variables in the current model will be based on both views. To explain the differences in regional unemployment and employment rates the following factors will be controlled for:

- Human capital variable: the share of the population aged 25–64 with a higher education (degree from university, higher technical institution, etc.) is included in the analysis (variable "Higher education"). For various reasons pointed out in the previous section, a higher share of high-skilled individuals increases the speed of adjustment in the labour market. Therefore, the variable is expected to have a negative relationship with unemployment rate and positive relationship with the employment rate.
- Demographic variables: to account for the age structure of the population, the share of young people (aged 15–24) in the working age population (aged 15–64) is included (variable "Youth"). As explained above, lower moving costs and lower risk aversion combined with barriers entering the labour market make the effect of the share of youth in the labour market ambiguous.

- Industrial composition: share of manufacturing and share of services in regional total employment are used as controls (variables "Manufacturing" and "Services"). As mentioned in the previous section, regions specialized in declining industries are assumed to exhibit higher unemployment rates than regions specialized in growing industries; however, empirical studies have shown mixed or not significant results in this aspect.
- Cross-country differences: to account for the cross-country differences in institutions and legislation between regions in different European countries, country dummies are added as control variables.

3.2 Spatial autocorrelation

Spatial dependence is accounted for using the spatial weight matrix *W*, which determines the structure and the intensity of spatial dependence between the regions. There are various ways to specify the spatial weights matrix and the specification of the matrix may influence the estimation results. The choice of a spatial weights matrix is somewhat arbitrary, as the structure of spatial interactions is not known a priori.

One of the simplest weight specifications frequently applied is a binary spatial weight matrix such that the elements of the matrix are $w_{ij} = 1$ if regions i and j share a border and otherwise $w_{ij} = 0$ (e.g. Diaz 2016). However, a simple binary matrix is not always appropriate as it assumes that spatial autocorrelation only occurs between the nearest neighbouring spatial units regardless of their size and shape (Cliff and Ord 1969, Getis 2009). Another group of spatial weights matrices are based on functions of the distances between the spatial units. One advantage of this type of matrix is that it allows all the weights between the regions to be positive and thus does not constrain the effective area (Cliff and Ord 1981).

As the regions in the current dataset are diverse in terms of size and shape and the aim is to also account for the spatial interaction between the regions that are not direct neighbours, the distance-based matrix is used.[4] The elements of the matrix used in the current analysis are constructed as the inverse values of distances between the physical geographic centres of the regions:

$$W = \begin{pmatrix} 0 & \frac{1}{d_{12}} & \cdots & \frac{1}{d_{1N}} \\ \frac{1}{d_{21}} & 0 & \cdots & \frac{1}{d_{2N}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{d_{N1}} & \frac{1}{d_{N2}} & \cdots & 0 \end{pmatrix}$$

The spatial weight matrix is row-standardized for easier interpretation following the common practice (e.g. Mitchell and Bill 2004, Niebuhr 2003, Semerikova 2015). Row-standardizing normalizes *W* so that the elements of each row sum to unity; in other words, the effect of the weighting operation can be interpreted as averaging over neighbouring values (Elhorst 2014).

[4] The author would like to thank the anonymous referee, who pointed to the possibility of robustness checks using different weight matrices. Unfortunately, data restrictions do not permit the author to use these specifications, such as a weight matrix based on cultural closeness or on distances between economic (instead of geographic) centres of regions. A weight matrix based on the GDP of the regions provides an interesting approach, indicating that migration is motivated not by physical distances, but rather by differences in the economic conditions of the regions. However, the aim of this paper is to focus both on migration and commuting behavior. While in terms of the latter geographical distance plays a key role, the use of a GDP-based weight matrix falls out of the scope of the current study.

The preliminary evidence of spatial clustering and potential spatial interaction is given by measures of spatial autocorrelation. In this case Moran's I is used.[5] The estimated Moran's I for the unemployment and employment rate are 0.18 and 0.262 (z-values are 33.44 and 48.18 respectively). Thus, there is a significant positive spatial correlation both for unemployment and employment rates. The positive sign of the spatial autocorrelation for the employment rate is a preliminary indicator that there are cooperation effects between industries in different regions. Negative dependence would have indicated that the competition for labour force between regions is stronger than the cooperation effects.

The presented Moran's I gives the value of the global spatial autocorrelation. To find out how spatial autocorrelation varies across regions, local measures of spatial autocorrelation are used. Local measures also allow us to single out the specific regions that exhibit significant autocorrelation and are therefore the ones potentially having the strongest interaction with their neighbours. To measure the local spatial autocorrelation, local indicators of spatial association (LISAs) for Moran's I are used.[6]

Figure 3 displays the areas with significant LISAs for unemployment rates. The darkest and lightest grey areas on the map exhibit significant positive autocorrelation. The darkest ones, marked by *high-high*, are regions with high unemployment surrounded by other regions with high unemployment. The lightest grey marks areas with low unemployment, whose neighbours also exhibit low unemployment. Areas marked by high-low (low-high) represent significant negative autocorrelation where high (low) unemployment areas are surrounded by low (high) unemployment areas. The spatial associations indicated in this map are in line with what was seen in Figure 1 as preliminary evidence of clustering. Clusters of high unemployment form in Spain and Portugal, southern Italy and Greece. Low unemployment clusters can be seen for regions in the UK and Norway. It is worth noting that there are only a few regions, namely in France, Belgium, Bulgaria and Turkey that show evidence of a negative association. While the clusters inside countries might be explained by some country level characteristics, perhaps the most interesting indication of the map is the positive spatial association of Germany with its southern, eastern and northern neighbours. This association is likely to be the result of commuting and migrating across national borders, especially between the areas that share a common language.

[5] To measure spatial autocorrelation, Moran's I statistic (Moran 1948) is calculated in the following way:

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}\,(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

where N is the number of regions indexed by *i* and *j*, x_i and x_j are the variables of interest in regions *i* and *j*, $\bar{x}$ is the average of x over N regions and w_{ij} is the element of spatial weight matrix *W* summarizing the interaction between regions *i* and *j*. N=306 in current case.

[6] Local indicators of spatial association (LISAs) for Moran's I (Anselin 1995) are defined as follows:

$$I_i = \frac{N(x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2} \sum_j w_{ij}(x_j - \bar{x})$$

where the notation is as described above. The spatial weight matrix W, as defined in the equation on page 13, is used to calculate LISAs.

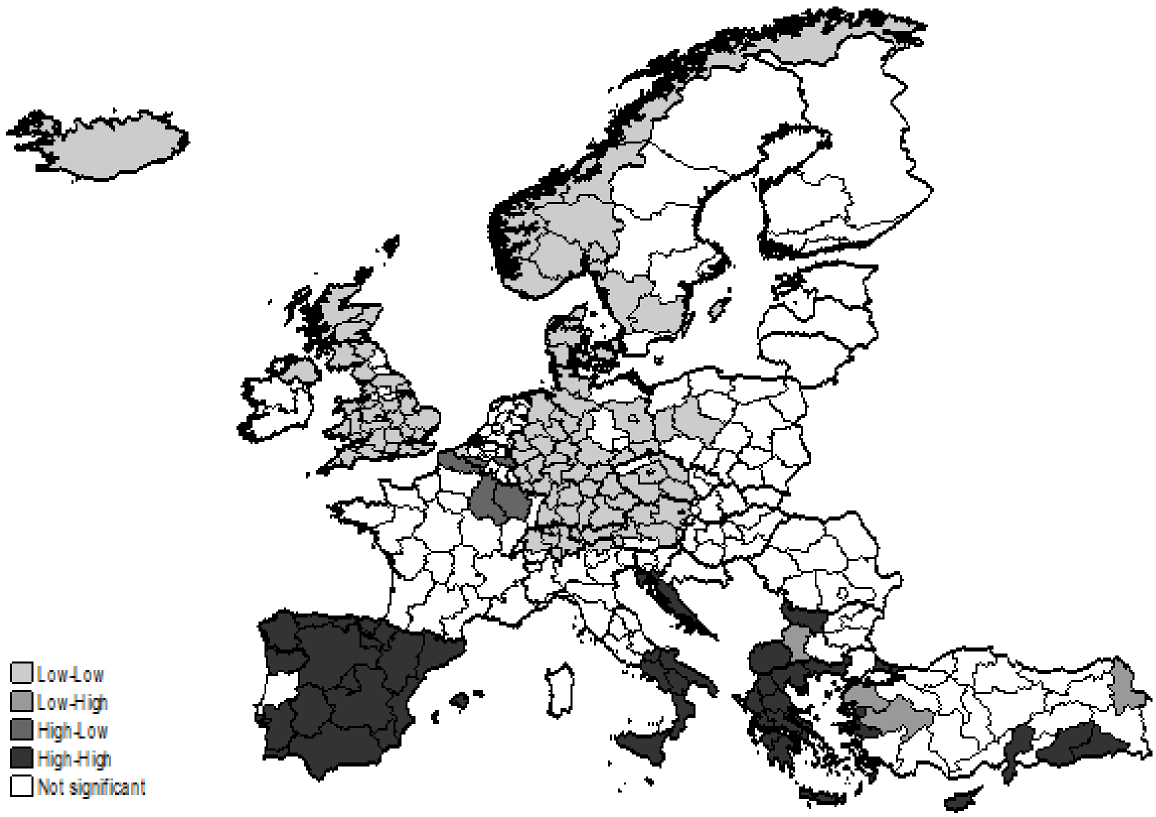

Figure 3. Local indicators of spatial association for unemployment rates (Eurostat database; compiled by the author)

Areas with significant LISAs for employment rates are shown in Figure 4. Again, clusters inside country borders are visible – in the UK, Norway and Sweden. Although spatial associations in southern Italy and Greece were also found to be significant in terms of unemployment rates, and here a bigger cluster is forming with regions in Greece, Turkey, Bulgaria and Romania, all exhibiting positive spatial autocorrelation with their neighbours. It should be noted that regions in Spain and Portugal are not found to be significant in terms of spatial association for employment rates. As in Figure 3, regions in Germany and its neighbours stand out as those with significant spatial association. In this case also regions in the Netherlands are part of the cluster. Only a few regions exhibit negative spatial associations. All of those regions are at international borders and the negative associations seen in the raw data could be therefore explained by cross-country differences in institutions and legislation.

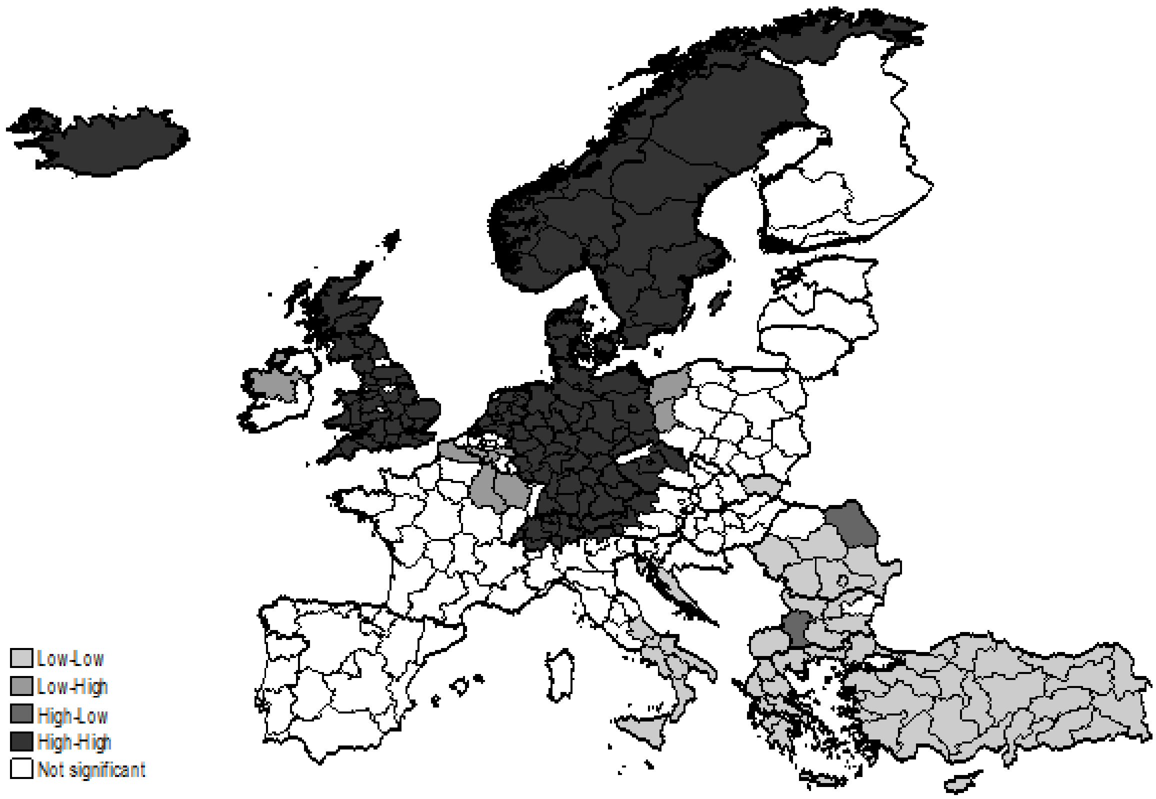

Figure 4. Local indicators of spatial association for employment rates (Eurostat database; compiled by the author)

The preliminary analysis so far indicates that both regional unemployment and employment rates in Europe exhibit positive spatial autocorrelation. Looking at the regions more closely, a few were found that also have negative associations with their neighbours, though a positive relationship clearly dominates. While some regions form clusters inside countries, data on regions in Germany, Switzerland, Austria, Denmark, Poland (for unemployment rates) and the Netherlands (for employment rates) point to an existing spatial association across national borders. To determine whether the spatial association can be explained due to similar regional characteristics or is partly a sign of cross-border interaction (e.g. in the form of commuting), a regression analysis is carried out.

3.3 Spatial models

In a spatial econometric model, three different types of spatial interaction effects can be distinguished: endogenous interaction effects among the dependent variable (Y), exogenous interaction effects among the independent variables (X) and interaction effects among the error terms (u).

Endogenous interaction effects result from direct interaction between regions and can be explained as part of an equilibrium outcome of a spatial interaction process. In this case the value of the dependent variable for one unit is jointly determined with that of neighbouring units. Endogenous interaction effects reflect the substantive form of spatial autocorrelation. In the case of an exogenous interaction effect, the dependent variable of a particular unit depends on the independent explanatory variables of other units. Interaction effects among the error terms reflect a situation where determinants of the dependent variable omitted from the model are spatially autocorrelated, a situation where unobserved shocks follow a spatial pattern or the case of measurement errors, where the regional system is wrongly specified and does not reflect the spatial structure of economic activities (Elhorst 2014). Interaction effects restricted to error terms account for a nuisance form of spatial dependence.

This study estimates the spatial error, spatial lag, spatial autoregressive model with spatial autoregressive disturbances and spatial Durbin model, each accounting for a different spatial interaction effect. The spatial lag model (SLM), also known as the spatial autoregressive model (SAR), accounts for the endogenous interaction effect, and the spatial error model (SEM) accounts for the interaction effect in error terms; both are presented in the seminal book by Anselin (1988). Ignoring spatially lagged dependent variables may lead to biased and inefficient estimates. Ignoring spatially correlated errors may result in inefficient estimates.

The spatial error model (SEM) takes the form:

$$\boldsymbol{y} = \alpha\boldsymbol{\iota}_N + \boldsymbol{X\beta} + \boldsymbol{u}$$

$$\boldsymbol{u} = \lambda\boldsymbol{Wu} + \varepsilon$$

where $\boldsymbol{y}$ is a $(N \times 1)$ vector of dependent variables, $\boldsymbol{X}$ is a $(N \times k)$ matrix of k explanatory variables, $\boldsymbol{\iota}_N$ is a $(N \times 1)$ vector of ones, α is a constant term parameter, $\boldsymbol{\beta}$ is a $(k \times 1)$ vector of parameters, λ is a scalar of the spatial autocorrelation coefficient restricted to interval (-1, 1), $\boldsymbol{W}$ is the $(N \times N)$ spatial weight matrix and random term $\varepsilon \sim N(0, \sigma^2)$.

The spatial lag model (SLM) is defined as:

$$\boldsymbol{y} = \rho\boldsymbol{Wy} + \alpha\boldsymbol{\iota}_N + \boldsymbol{X\beta} + \varepsilon$$

where ρ is the spatial autoregressive coefficient, restricted to interval (-1, 1) and the other notation is as described above.

In addition to the SLM and SEM model, a spatial autoregressive model with spatial autoregressive disturbances (SARAR model) is also estimated (see Kelejian and Prucha 1998). The SARAR model incorporates both the endogenous interaction effects and the interaction effects among error terms:

$$\boldsymbol{y} = \rho\boldsymbol{Wy} + \alpha\boldsymbol{\iota}_N + \boldsymbol{X\beta} + \boldsymbol{u}$$

$$\boldsymbol{u} = \lambda\boldsymbol{Wu} + \varepsilon$$

where the notion is as described above.

To account for the spatial effects among independent variables, the spatial Durbin model is estimated. The spatial Durbin model (SDM) (see LeSage and Pace 2009) takes the form:

$$\boldsymbol{y} = \rho\boldsymbol{Wy} + \alpha\boldsymbol{\iota}_N + \boldsymbol{X\beta} + \boldsymbol{WX\theta} + \varepsilon$$

where $\boldsymbol{\theta}$ is a $(N \times 1)$ vector of coefficient estimates of the spatially lagged values of explanatory variables and the other notation is as described above.

All spatial models are estimated with the maximum likelihood estimation. The estimation was carried out using Stata software. To calculate direct and indirect effects Matlab routines developed by LeSage (1999) were also used.

3.4 Direct and indirect effects

While changes in explanatory variables in region *i* are likely to affect the (un)employment rate in the same region, the effect of the changes on (un)employment rates in other regions is also of interest in the present study. To investigate those effects, summary measures of direct and indirect spatial effects are estimated following the methodology proposed by LeSage and Pace (2009).

It is possible to rewrite the spatial lag model as:

$$(I - \rho W)y = X\beta + \boldsymbol{\iota}_N \alpha + \varepsilon$$

$$y = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1} \boldsymbol{\iota}_N \alpha + (I - \rho W)^{-1} \varepsilon$$

Following LeSage and Pace (2009), the matrix $(I - \rho W)^{-1} I_N \beta_r$ will be denoted as $S_r(W)$. The derivative of y_i with respect to x_{ir} is represented by the *ii*-th element of the matrix $S_r(W)$ and denotes the direct effect:

$$\frac{\partial y_i}{\partial x_{ir}} = S_r(W)_{ii}$$

The average direct effect is the average of the diagonal elements of the matrix $S_r(W)$. It measures the summary impact of the changes of variable r in the *i*th region on the dependent variable in the same region using an average over the regions. The derivative of y_i with respect to x_{jr} is represented by the ij-th element of the matrix $S_r(W)$ and denotes the indirect (i.e. spillover) effect:

$$\frac{\partial y_i}{\partial x_{jr}} = S_r(W)_{ij}$$

The average indirect effect is the average of row-sums of the non-diagonal elements of the matrix $S_r(W)$. This measure reflects the impact of changes of variable r in one region on the value of the dependent variable in all other regions. It is important to note that the average indirect effect measures cumulative impacts over all regions, and therefore often exceeds the average direct effect.

4. RESULTS

Table 1 reports the estimates from the OLS, SEM, SLM and SARAR models, explaining the regional unemployment rate differentials.[7] As mentioned above, in the presence of spatial interactions in the model, OLS estimates are biased and/or inefficient. To test for the spatial dependence in OLS residuals, Moran's I could be used. However, Moran's I does not identify whether spatial autocorrelation results from the endogenous interaction effects or interaction effects among the error terms. The Lagrange multiplier (LM) tests, presented by Anselin *et al.* (1996), are used to test for both types of interaction.[8] The results from the LM-tests (see Table 1) reject the null of no spatial dependence against both forms of spatial dependence. The results from the robust LM-test indicate that the SLM model would be the best specification compared to the SEM model. Based on the Akaike information criterion (AIC), the SARAR model seems to be the best choice; however, Bayesian information criterion (BIC) is marginally lower for the SLM model.

The estimates show interesting features – most of them consistent with the theoretical background and empirical literature that analyses unemployment rates in different regions and countries, presented in the previous section. It should be noted that the coefficient estimates differ only marginally between the different models. All of the estimates are statistically significant, except the share of manufacturing. A higher share of young people is significantly related to a higher regional unemployment rate. This finding is in line with the results from López-Bazo *et al.* (2002), Mitchell and Bill (2004) and Semerikova (2015). The positive relationship indicates the barriers the younger generation has in terms of entering the labour market.

Table 1. OLS, SEM, SLM and SARAR estimates of regional unemployment rate determinants

Unemployment	OLS	SEM	SLM	SARAR
Youth	0.329**	0.356**	0.340**	0.363**
	[0.074]	[0.071]	[0.066]	[0.068]
Services	0.155**	0.158**	0.165**	0.168**
	[0.033]	[0.030]	[0.029]	[0.029]
Manufacturing	-0.047	-0.036	-0.033	-0.023
	[0.036]	[0.034]	[0.032]	[0.032]
Higher education	-0.165**	-0.172**	-0.167**	-0.176**

[7] Results for the SDM model are reported in appendix 3. Almost all of the spatial lags of independent variables are not statistically significant for unemployment and employment rate models, meaning that the SDM model does not provide additional information compared to the SLM model. Therefore, SDM results are not commented further.

[8] According to the decision rule, spatial dependence is of the spatial lag form if the LM-test for spatial lag dependence (LM-Lag) is more significant than the test for spatial error dependence (LM-Error) and the robust version of LM-Lag, which is robust against the presence of spatial error dependence, is significant. The opposite indicates that the spatial dependence is of the spatial error form.

	[0.034]	[0.032]	[0.030]	[0.031]
Constant	-0.059	-0.061	-0.135**	-0.139**
	[0.033]	[0.033]	[0.030]	[0.030]
λ		0.883**		0.813**
		[0.120]		[0.173]
ρ			0.929**	0.918**
			[0.069]	[0.078]
Country dummies	YES	YES	YES	YES
AIC	-1308.08	-1313.13	-1332.96	-1335.98
BIC	-1170.30	-1167.91	-1187.74	-1187.04
R^2	0.808			
LM-Error	7.42**			
Robust LM-Error	2.92			
LM-Lag	38.47**			
Robust LM-Lag	33.97**			

Note: * significant at 5%; ** significant at 1%. Standard errors are in brackets. N=306

The results from the estimates for the industrial composition of the regional labour market seems at first somewhat unexpected. According to the theory, one would expect regions specialized in declining industries such as agriculture and manufacturing to exhibit higher unemployment rates than regions specialized in growing industries. However, no significant results for the share of manufacturing employment and a positive relationship for the share of services appear. Empirical evidence from earlier studies also showed mixed or not significant results (see e.g. Semerikova 2015, López-Hernández 2013, Filiztekin 2009, Diaz 2016). The results could be explained by the fact that in areas with higher unemployment rates the former unemployed have found employment opportunities in the services sector. As an individual, it is easier to create opportunities for services (e.g. through self-employment in home accommodation etc.) than to develop large scale manufacturing that needs more time and investment. Therefore, the short-term services sector is more flexible in responding to the movements of unemployment than the manufacturing sector. Another explanation for the results can be that services and manufacturing sectors are internally heterogeneous; that is, the manufacturing and services sectors include a wide range of different skill level jobs. This opportunity will be investigated below.

The share of the population with a higher education has an expected relationship with the unemployment rate. In regions with a higher share of high-skilled people, the unemployment rate is lower. This is in line with earlier empirical results (e.g. Diaz 2016, López-Bazo *et al.* 2002, López-Hernández 2013) and also with the theory that points to the faster speed of adjustment of highly educated people. Most of the coefficients for country dummies are found to be significant in the models included here, which indicates that cross-country differences in institutions and legislation is an important factor in terms of unemployment rate differences between regions.

Perhaps the most interesting elements of the results are the spatial autocorrelation and spatial autoregressive coefficient estimates that provide an answer to the earlier question about the existence and statistical significance of spatial interactions between regional labour markets. While the preliminary analysis carried out in the previous section did point to significant spatial associations among regional labour markets in Europe, it was not clear if the significant relationship only reflected the spatial clustering of regions with similar characteristics or pointed to the effects of spatial interaction (e.g. in the form of commuting). The models presented here control for age structure, human capital, industrial structure and the role of institutions and still find the spatial autocorrelation coefficient λ in the SEM model and the spatial autoregressive coefficient ρ in the SLM model to be significant. The spatial coefficient estimates in the SARAR model are also both significant and similar to the respective coefficient estimates in the SEM and SLM models. It is worth noting that the significance of both spatial coefficients shows that both the substantive and nuisance forms of spatial autocorrelation exist. Therefore, the unemployment rate in one region is directly affected by unemployment rate changes in other regions, but also by unobserved shocks in other regions. Overall, the results point to the fact that there exist spillovers across regional labour markets, which can be expressed, for example, in the form of workers commuting from one region to another.

The estimates from different models for the employment rate are presented in Table 2. As for the unemployment rate, the LM-tests reject the null of no spatial dependence in OLS model residuals. The LM-tests indicate that the SLM would be the best model compared to the SEM model. Based on the information criteria, the most general SARAR model is the preferred specification. As in the case of unemployment rates, the coefficient estimates are similar across all models. All of the estimates are statistically significant, except the share for manufacturing. As expected, areas with a higher share of young population have on average lower employment rates. The results in terms of industrial composition are in line with the results in the case of unemployment rates. The share for manufacturing is also not found to be significant here and the share for services has a negative relationship with regional employment rates. The positive relationship with the higher share of higher educated is as expected and reflects the fact that more educated people are more efficient at finding jobs and also more demanded in the labour market. Most of the country dummies are found to be statistically significant, reflecting the fact that cross-country differences in institutions is an important determinant of regional labour market differences.

Table 2. OLS, SEM, SLM and SARAR estimates of regional employment rate determinants

Employment	OLS	SEM	SLM	SARAR
Youth	-0.733**	-0.753**	-0.689**	-0.708**
	[0.108]	[0.103]	[0.095]	[0.097]
Services	-0.244**	-0.249**	-0.259**	-0.263**
	[0.048]	[0.044]	[0.042]	[0.042]

Manufacturing	0.045	0.024	0.005	-0.012
	[0.053]	[0.049]	[0.047]	[0.046]
Higher education	0.314**	0.330**	0.312**	0.332**
	[0.050]	[0.046]	[0.043]	[0.044]
Constant	0.926**	0.925**	0.279**	0.289**
	[0.048]	[0.052]	[0.058]	[0.063]
λ		0.921**		0.891**
		[0.080]		[0.083]
ρ			0.941**	0.933**
			[0.058]	[0.065]
Country dummies	YES	YES	YES	YES
AIC	-1073.79	-1086.52	-1106.57	-1116.62
BIC	-936.01	-941.29	-961.35	-967.67
R^2	0.8509			
LM-Error	22.09**			
Robust LM-Error	0.62			
LM-Lag	48.24**			
Robust LM-Lag	26.77**			

Note: * significant at 5%; ** significant at 1%. Standard errors are in brackets. N=306

The results also show spatial interactions in the case of the employment rate. The spatial autocorrelation coefficient in the SEM model and the spatial autoregressive coefficient in the SLM model are both significant and of high value. Furthermore, in the SARAR model both estimates of the spatial coefficient are significant and their values are respectively similar to those in the case of the SEM and the SLM model. Therefore, both substantive and nuisance forms of spatial autocorrelation exist for employment rates. While positive spatial dependence could not be confirmed in terms of employment rates based on earlier studies (e.g. Pavlyuk 2011, Lewis *et al.* 2011, Mayor and López 2008), the results of this study provide support for the existence of positive spillovers also in the case of employment. The results indicate that instead of competition for labour force between regions, which would result in negative dependence, cooperation effects dominate here, resulting in positive dependence.

To investigate the nature of the spatial spillover further, direct and indirect effects are estimated based on the SLM model. The estimates for unemployment rates are given in Table 3. The direct effects estimates are similar to the coefficient estimates in the SLM model (see Table 1). As before, the coefficient of the share for manufacturing is not found to be significant. The direct effects estimates for other variables lead to the same conclusions as the coefficient estimates for the SLM model.

The average indirect effect shows the effect that changes have in each explanatory variable in one region on unemployment in other regions. It is important to note that this measure shows a cumulative impact over space. Therefore, the indirect effect is often estimated to be higher than the direct effect. Interestingly, indirect effects are not found to be significant in this case. Therefore, no evidence is found that the spatial effects work through the differences in the demographics, such as the population of youth, differences in industrial structure or differences in human capital.

Table 3. Direct and indirect effect estimates for unemployment rate (based on SLM)

Unemployment	Direct effect	Indirect effect
Youth	0.354**	8.594
	[5.02]	[1.23]
Services	0.176**	4.308
	[5.49]	[1.21]
Manufacturing	-0.032	-0.779
	[-0.94]	[-0.63]
Higher education	-0.178**	-4.339
	[-5.60]	[-1.24]

Note: * significant at 5%; ** significant at 1%. t-statistics are in brackets. N=306

The direct and indirect estimates of employment rates are presented in Table 4. Direct effects are again slightly higher than the coefficient estimates of the SLM model (see Table 2). The indirect effects are not found to be statistically significant for employment rates. Therefore, spatial effects do not seem to work either for employment rates through the differences in the demographics, differences in industrial structure or differences in human capital. Again, the exact mechanism of spatial dependence remains an interesting question for future research.

Table 4. Direct and indirect effect estimates for employment rate (based on SLM)

Employment	Direct effect	Indirect effect
Youth	-0.756**	14.413
	[-6.39]	[-1.13]
Services	-0.282**	-5.440
	[-5.60]	[-1.09]
Manufacturing	0.006	-1.095
	[0.12]	[0.08]
Higher education	0.342**	6.606
	[6.21]	[1.10]

Note: * significant at 5%; ** significant at 1%. t-statistics are in brackets. N=306

To investigate how spatial dependence has involved in Europe since the Eastern enlargement of the European Union in 2004, SEM, SLM and SARAR models have been estimated for the years 2004, 2008, 2011 and 2015. To aid the comparability of the results, the sample of all years includes data on 253 NUTS 2 regions. The spatial coefficients of the estimated models for unemployment rates are presented in Table 5. First, spatial dependence in all the given years is rather similar. By comparing 2004 and 2015, it can be seen that spatial dependence is slightly higher. However, the years in between exhibit a slightly lower level of dependence. With the presence of spatial lag dependence, spatial error dependence turns out to be significant only for the most recent year.

Table 5. Spatial coefficients for 2004, 2008, 2011 and 2015 for the unemployment rate

Unemployment		2004	2008	2011	2015
SEM	λ	0.803**	0.794**	0.742**	0.856**
SLM	ρ	0.907**	0.863**	0.851**	0.926**
SARAR	λ	0.64	0.654	0.369	0.720**
	ρ	0.895**	0.835**	0.836**	0.913**

Note: * significant at 5%; ** significant at 1%. N=253

The spatial coefficients of the estimated models for employment rates are also rather stable over time (see Table 6). Spatial dependence has only raised slightly over the years. The spatial error dependence remains significant in the given models even with the presence of spatial lag dependence.

Table 6. Spatial coefficients for years 2004, 2008, 2011 and 2015 for the employment rate

Employment		2004	2008	2011	2015
SEM	λ	0.863**	0.854**	0.884**	0.897**
SLM	ρ	0.912**	0.927**	0.949**	0.947**
SARAR	λ	0.777**	0.698*	0.725*	0.802**
	ρ	0.897**	0.914**	0.940**	0.937**

Note: * significant at 5%; ** significant at 1%. N=253.

These results may seem somewhat puzzling at first, as one might expect the spatial dependence, as a sign of European labour market integration, to grow considerably over the years. The slight decrease in 2008 could be related to the economic crisis. In addition, factors that lessen the need to commute or migrate for work to neighbouring regions could presumably prevent the spatial dependence from growing over the period. One of the reasons could be the enhanced possibilities for remote working. Modern communication possibilities enable working remotely without having the need to migrate. Another possible explanation is that the difference in average wages and overall living standards between the different European regions have decreased compared to the time of the Eastern enlargement of EU. Therefore, the gain from migrating or commuting for work has also diminished.

5. ROBUSTNESS CHECKS

To check for the robustness of the results over time, panel data models were used. Data for 2004, 2008, 2011 and 2015 for 253 NUTS 2 regions were used.[9] The pooled data (OLS) model, SEM and SLM models with random effects were estimated. While country dummies are included in the models, fixed effects models are not considered because fixed effects models do not allow us to investigate the impact of time invariant explanatory variables. The SARAR model is excluded, while the random effects variant of the SARAR model can be written as a special case of the SLM specification (Belotti *et al.* 2016).

The results of the panel data models on the unemployment rate are presented in Appendix 4. All of the coefficient estimates, and the share of manufacturing, are significant in these models. While the exact values of the coefficients are somewhat different than for earlier results, the signs of the coefficients are in accordance with the earlier results (see Table 1). Overall, the coefficient estimates lead us to the same conclusions as the results on cross-sectional data. In particular, the spatial autocorrelation and spatial autoregressive coefficient are also positive, significant and of slightly higher value as in the cross-sectional data. It can be concluded that even after taking the time dimension into account, the spatial dependence of both nuisance and substantive forms remains.

Appendix 5 displays the results of the panel data model on employment rates. The share of manufacturing is found to be significant here in the pooled data model. The share of people with

[9] Some of the initial 306 regions were excluded because of data unavailability for 2004, 2008 and 2011.

higher education is not found to be statistically significant in the SLM model. Overall, the values of the coefficients are similar to the ones obtained above and lead us to similar conclusions about the relationship between the changes in the explanatory variables and the employment rate (see Table 2). The spatial autocorrelation and spatial autoregressive coefficient are also positive, significant and of slightly higher value as in the cross-sectional data. Therefore, the spatial dependence remains also for regional employment rates after accounting for the time dimension.

While the results from the estimates for industrial composition were somewhat unexpected, the robustness of the results is checked using the subcategories of the sectors. The manufacturing sector includes jobs covering a wide range of different skill levels, which we have tried to consider by dividing the share of employment in manufacturing into two categories: high-technology manufacturing (HTM) and low-technology manufacturing (LTM). The services sector is also divided into two subcategories: knowledge-intensive services (KIS) and less knowledge-intensive services (LKIS). The results for unemployment and employment rates are presented in Appendix 6 and Appendix 7. In the case of unemployment rates, dividing services into two subsectors does not alter the results, both KIS and LKIS have positive and significant estimates. The share of high-technology manufacturing is statistically significant in three out of four models. The share of low-technology manufacturing remains insignificant. These findings again seem to point to the fact that services and here also the high-technology manufacturing sector are more capable of adapting to changes in unemployment than low-technology manufacturing. In the case of the employment rate (Appendix 7), dividing manufacturing into subcategories does not have any effect; the estimates of HTM and LTM are both found to be insignificant in all models. Estimates of shares for both subcategories of services remain negative and significant, except for LKIS in two of the models. Other parameter estimates do not substantially differ from the earlier results presented in Table 2.

6. CONCLUSIONS

The results of this study emphasize the importance of spatial interaction in regard to regional labour markets in Europe. The spatial dependence of unemployment and employment rates is investigated using data on 306 NUTS2 regions in Europe. The findings show that regional labour markets in Europe cluster in space; that is, regions with high (low) unemployment/employment rate are surrounded by regions with high (low) unemployment/employment rate.

Spatial dependence is explored using different types of spatial econometrics models that account for spatial effects working through a dependent variable, independent variables and an error term. The results are stable throughout the models. The set of factors is controlled to determine whether the spatial association is explained by the clustering of regions with similar characteristics. The factors affecting regional labour markets are chosen based on the equilibrium and disequilibrium view. The findings of this study suggest that differences across regions in age structure, sector specialization, human capital and country level institutions are factors behind observed unemployment and employment rate disparities. However, even after controlling for these factors, spatial dependence remains significant for unemployment and employment rates. Spatial dependence is found to be positive both for unemployment and employment rates, indicating that cooperation effects between regions dominate over competition for labour force effects. Both substantive and nuisance forms of spatial dependence exist; that is, the (un)employment rate in one region is directly affected by (un)employment rate changes in other regions, but also by unobserved shocks in other regions. The results point to the fact that there exist significant spillovers across the regional labour markets. Interestingly, spatial dependence between regional labour markets in Europe has been fairly stable throughout the years starting from the Eastern enlargement of the European Union in 2004. No evidence is found that the spatial effects work through the changes in the demographics, such as the share of the population of youth, changes in industrial structure or changes in human capital. The exact mechanism of spatial dependence remains an interesting question for future research.

Another challenge of future research would be estimating the models based on NUTS 3 data. Using NUTS 3 regions would provide data on smaller units and is likely to lead to more significant results as it would allow researchers to also capture the interactions between regions inside the borders of smaller countries. For example, instead of one region at NUTS 2 level, Estonia is divided into 5 regions at NUTS 3 level. Unfortunately, to date the Eurostat database provides regional labour market data only down to NUTS 2 level. Providing data on smaller regional units would open up interesting research possibilities.

The findings of this study provide information for regional and labour market policy measures in Europe. Policy measures aiming to reduce regional unemployment and enhance employment should take into account the spatial interaction between the labour markets and should be therefore coordinated between the neighbouring regions. For example, enhancing infrastructure and transport connections between neighbouring regions could benefit both regions as it allows for an increased commuting of workers, and therefore more efficient matching of workers to jobs based on skills. All in all, reducing labour market differences calls for close cooperation between regions.

REFERENCES

Anselin L., Bera A.K., Florax R., Yoon M.J. (1996) "Simple diagnostic tests for spatial dependence." *Regional Science and Urban Economics*, vol. 26, no. 1, pp. 77–104.

Anselin, L. (1988) "Spatial Econometrics: Methods and Models." Kluwer Academic Publishers, Dordrecht.

Anselin, L. (1995) "Local indicators of spatial association – LISA." *Geographical Analysis*, vol. 27, no. 2, pp. 93–115.

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-Pyrénées region." *Papers in Regional Science,* vol. 82, pp. 155–174.

Badinger, H., Url, T. (2002) "Determinants of regional unemployment: some evidence from Austria." *Regional Studies,* vol. 36, no. 9, pp. 977–988.

Belotti, F., Hughes, G. Mortari, A.P. (2016). "Spatial panel models using Stata." CEIS Tor Vergata, Research Paper Series, vol. 14, issue 5, no. 373

Beyer, R.C.M., Stemmer, M.A. (2016) "Polarization or convergence? An analysis of regional unemployment disparities in Europe over time." *Economic Modelling*, vol. 55, pp. 373–381.

Cliff, A.D., and J.K. Ord. (1969) "The problem of spatial autocorrelation." *In London Papers in Regional Science*, A. J. Scott (ed.), Pion, London, pp. 25–55.

Cliff, A.D., Ord, J.K. (1981) "Spatial Processes: Models and Applications." Pion, London.

Cracolici, M. F., Cuffaro, M., Nijkamp, P. (2007) "Geographical distribution of unemployment: An analysis of provincial differences in Italy." *Growth and Change, vol.* 38, no. 4, pp. 649–670.

Diaz, A.M. (2016) "Spatial Unemployment Differentials in Colombia." *Desarrollo y Sociedad,* issue 76, pp. 123–163.

Elhorst, (2003) "The mystery of regional unemployment differentials: Theoretical and empirical explanations." *Journal of Economic Surveys,* vol. 17, no. 5, pp. 709–748.

Elhorst, J.P. (2014) "Spatial Econometrics: From Cross-sectional Data to Spatial Panels." Springer, Berlin.

Eurostat database. European Commission. [http://ec.europa.eu/eurostat/data/database] 28/02/2017

Filiztekin A. (2009) "Regional unemployment in Turkey." *Papers in Regional Science*, vol. 88, no. 4, pp. 863–878.

Getis, A. (2009) "Spatial Weights Matrices." *Geographical Analysis*, vol. 41, no. 4, pp. 404–410.

Kelejian, H.H., Prucha, I.R. (1998) "A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances." *Journal of Real Estate Finance & Economics*, vol. 17, no. 1, pp. 99–121.

Kondo, K. (2015) "Spatial persistence of Japanese unemployment rates." *Japan and the World Economy*, vol. 36, pp. 113–122.

LeSage, J.P. (1999) "Econometrics Toolbox." [http://www.spatial-econometrics.com/] 20.04.2017

LeSage, J.P., Pace, R.K. (2009) "Introduction to spatial econometrics." Chapman and Hall/CRC, Boca Raton, Florida.

Lewis Jr., W., DiFurio, F., Goode, T. (2011) "Testing for the Presence of Spatial Dependence in Manufacturing Employment in South Carolina Counties." *Southern Business & Economic Journal*, vol. 34, no. 1/2, pp. 41–53.

Lopez-Bazo, E., Del Barrio, T., Artis, M. (2002) "The regional distribution of Spanish unemployment: A spatial analysis." *Papers in Regional Science,* vol. 87, no. 3, pp. 365–389.

López-Hernández, F. A. (2013) "Second-order polynomial spatial error model. Global and local spatial dependence in unemployment in Andalusia." *Economic Modelling*, vol. 33, pp. 270–279.

Marston, S.T. (1985) "Two views of the geographic distribution of unemployment." *Quarterly Journal of Economic*, vol.100, no. 1, pp. 57–79.

Mayor, M., López, A.J. (2008) "Spatial shift-share analysis versus spatial filtering: an application to Spanish employment data." *Empirical Economics*, vol. 34, pp. 123–142.

Mitchell, W., Bill, A. (2004) "Spatial dependence in regional unemployment in Australia." Centre of Full Employment and Equity The University of Newcastle, Working Paper No. 04-11, pp. 1–28.

Molho, I. (1995) "Spatial autocorrelation in British unemployment." *Journal of Regional Science*, vol. 35, no. 4, pp. 641–658.

Moran, P. (1948) "The interpretation of statistical maps." *Journal of the Royal Statistical Society . Series B,* vol. 10, no. 2, pp. 243–251.

Niebuhr, A. (2003) "Spatial interaction and regional unemployment in Europe." *European Journal of Spatial Development*, vol. 5, pp. 1–26.

NUTS overview. European Commission. [http://ec.europa.eu/eurostat/web/nuts/overview] 8/03/2017

Overman, H.G., Puga, D. (2002) "Unemployment clusters across Europe"s regions and countries." *Economic policy*, vol. 17, no. 34, pp. 117–147.

Patacchini, E., Zenou, Y. (2007). "Spatial dependence in local unemployment rates." *Journal of Economic Geography, vol.* 7, no. 2, pp. 169–191.

Pavlyuk, D. (2011) "Spatial Analysis of Regional Employment Rates in Latvia.", *Sustainable Spatial Development*, vol. 2, pp. 56–62.

Semerikova, E. (2015) "Spatial Patterns of German Labour Market: Panel Data Analysis of Regional Unemployment." Geographical Labour Market Imbalances: Recent Explanations and Cures, Springer, Berlin-Heidelberg, chapter 3, pp. 37–64.

Taylor, J. (1996) "Regional Problems and Policies: A European Perspective." *Australasian Journal of Regional Studies*, vol. 2, pp. 103–131.

APPENDICES

Appendix 1. Description of variables and used datasets

Variable	Description	Dataset in Eurostat database
Unemployment rate	Unemployed persons as a percentage of the economically active population (i.e. labour force or sum of employed and unemployed).	Unemployment rate by NUTS 2 regions [tgs00010]
Employment rate	Employed persons aged 15–64 as a percentage of the population of the same age group.	Employment rate of the age group 15–64 by NUTS 2 regions [tgs00007]
Higher education	Share of population aged 25–64 who have successfully completed tertiary studies (e.g. university, higher technical institution, etc.).	Tertiary educational attainment, age group 25–64 by sex and NUTS 2 regions [tgs00109]
Youth	Share of youth (aged 15–24) in the whole working age population (aged 15–64).	Population aged 15 and over by sex, age and NUTS 2 regions (1 000) [lfst_r_lfsd2pop]
Manufacturing	Share of manufacturing in regional total employment.	Employment in technology and knowledge-intensive sectors by NUTS 2 regions and sex (from 2008 onwards, NACE Rev. 2) [htec_emp_reg2]
Services	Share of services in regional total employment.	
KIS	Share of knowledge-intensive services in total employment	
LKIS	Share of less knowledge-intensive services in total employment	
HTM	Share of high-technology manufacturing in total employment	
LTM	Share of low-technology manufacturing in total employment	

Appendix 2. Countries included in the sample

Austria(9), Belgium(11), Bulgaria(6), Switzerland(7), Cyprus(1), Czech Republic(8), Germany(38), Denmark(5), Estonia(1), Greece(13), Spain(16), Finland(4), France(21), Croatia(2), Hungary(7), Ireland(2), Island(1), Italy(21), Lithuania(1), Luxembourg(1), Latvia(1), Macedonia(1), Malta(1), Netherlands(12), Norway(7), Poland(16), Portugal(5), Romania(8), Sweden(8), Slovenia(2), Slovakia(4), Turkey(26), United Kingdom(40)

Note: number of regions included per country in brackets.

Appendix 3. SDM models for unemployment and employment rate

	Unemployment	Employment
Youth	0.488**	-0.785**
	[0.068]	[0.097]
Services	0.172**	-0.268**
	[0.028]	[0.040]
Manufacturing	0.017	-0.099*
	[0.031]	[0.044]
Higher education	-0.172**	0.300**
	[0.027]	[0.039]
constant	0.39	2.283
	[0.795]	[1.297]
ρ	-1.645*	-0.429
	[0.758]	[0.620]
W*Youth	-0.072	-3.325
	[1.287]	[1.889]
W*Services	-0.508	0.254
	[0.711]	[1.014]
W*Manufacturing	-1.374	-0.155
	[0.733]	[1.044]
W*Higher education	0.355	-1.479*
	[0.412]	[0.584]
Country dummies	YES	YES
AIC	-1417.42	-1200.75
BIC	-1138.15	-921.48
R^2	0.8921	0.9226

Note: * significant at 5%; ** significant at 1%. Standard errors are in brackets. N=306.

Appendix 4. Panel data models for unemployment rate

Unemployment	Pooled	SEM_{re}	SLM_{re}
Youth	0.436**	0.577**	0.410**
	[0.069]	[0.067]	[0.058]
Services	0.090**	0.121**	0.144**
	[0.028]	[0.030]	[0.031]
Manufacturing	-0.151**	-0.107**	-0.087*
	[0.032]	[0.035]	[0.035]
Higher education	-0.099**	-0.143**	-0.063**
	[0.024]	[0.027]	[0.023]
constant	-0.045	-0.071	-0.168**
	[0.029]	[0.051]	[0.031]
λ		0.980**	
		[0.010]	
ρ			0.979**
			[0.010]
Country dummies	YES	YES	YES
AIC	-3818.69	-4254.96	-4252.09
BIC	-3661.26	-4082.77	-4079.90
R^2	0.479		

Note: * significant at 5%; ** significant at 1%. Standard errors are in brackets. N=1012.

Appendix 5. Panel data models for employment rate

Employment	Pooled	SEM_{re}	SLM_{re}
Youth	-0.860**	-0.613**	-0.452**
	[0.080]	[0.073]	[0.060]
Services	-0.198**	-0.204**	-0.239**
	[0.032]	[0.039]	[0.038]
Manufacturing	0.123**	0.017	-0.011
	[0.037]	[0.050]	[0.047]
Higher education	0.331**	0.168**	0.049
	[0.028]	[0.037]	[0.026]
constant	0.911**	0.929**	0.310**
	[0.034]	[0.045]	[0.040]
λ		0.966**	
		[0.017]	
ρ			0.965**
			[0.017]
Country dummies	YES	YES	YES
AIC	-3522.35	-4122.35	-4144.34
BIC	-3364.92	-3950.16	-3972.15
R^2	0.7494		

Note: * significant at 5%; ** significant at 1%. Standard errors are in brackets. N=1012.

Appendix 6. OLS, SEM, SLM and SARAR estimates of regional unemployment rate determinants (subcategories for manufacturing and services)

Unemployment	OLS	SEM	SLM	SARAR
Youth	0.380**	0.411**	0.376**	0.405**
	[0.073]	[0.069]	[0.066]	[0.068]
KIS	0.238**	0.230**	0.243**	0.234**
	[0.051]	[0.047]	[0.046]	[0.045]
LKIS	0.146**	0.154**	0.155**	0.164**
	[0.044]	[0.041]	[0.040]	[0.040]
HTM	-0.150*	-0.135*	-0.121*	-0.109
	[0.067]	[0.061]	[0.060]	[0.060]
LTM	-0.008	-0.006	-0.002	-0.001
	[0.053]	[0.049]	[0.048]	[0.047]
Higher education	-0.222**	-0.225**	-0.222**	-0.228**
	[0.039]	[0.036]	[0.035]	[0.035]
constant	-0.081**	-0.082**	-0.147**	-0.148**
	[0.030]	[0.029]	[0.028]	[0.029]
λ		0.871**		0.820**
		[0.131]		[0.170]
ρ			0.874**	0.849**
			[0.117]	[0.132]
Country dummies	YES	YES	YES	YES
AIC	-1329.58	-1334.38	-1342.03	-1345.82
BIC	-1185.78	-1183.22	-1190.86	-1190.97
R^2	0.8293			
LM-Error	8.23**			
Robust LM-Error	0.00			
LM-Lag	18.52**			

Robust LM-Lag	10.29**

Note: * significant at 5%; ** significant at 1%. Standard errors are in brackets. N=295.

Appendix 7. OLS, SEM, SLM and SARAR estimates of regional employment rate determinants (subcategories for manufacturing and services)

Employment	OLS	SEM	SLM	SARAR
Youth	-0.733**	-0.750**	-0.672**	-0.692**
	[0.112]	[0.105]	[0.098]	[0.100]
KIS	-0.485**	-0.472**	-0.507**	-0.492**
	[0.078]	[0.071]	[0.069]	[0.067]
LKIS	-0.101	-0.118	-0.124*	-0.141*
	[0.068]	[0.062]	[0.060]	[0.059]
HTM	0.15	0.112	0.085	0.054
	[0.102]	[0.093]	[0.090]	[0.088]
LTM	-0.012	-0.013	-0.052	-0.047
	[0.081]	[0.074]	[0.071]	[0.070]
Higher education	0.457**	0.465**	0.458**	0.470**
	[0.060]	[0.054]	[0.052]	[0.052]
constant	0.939**	0.937**	0.300**	0.309**
	[0.046]	[0.047]	[0.061]	[0.067]
λ		0.903**		0.864**
		[0.097]		[0.119]
ρ			0.930**	0.920**
			[0.068]	[0.077]
Country dummies	YES	YES	YES	YES
AIC	-1077.48	-1087.14	-1105.6	-1112.96
BIC	-933.69	-935.97	-954.44	-958.11

R^2	0.8682
LM-Error	17.08**
Robust LM-Error	0.31
LM-Lag	40.30**
Robust LM-Lag	23.53**

Note: * significant at 5%; ** significant at 1%. Standard errors are in brackets. N=295.

KOKKUVÕTE

Ruumilised seosed regionaalsetel tööturgudel Euroopas

Uuringu tulemused rõhutavad ruumiliste seoste tähtsust Euroopa regionaalsetel tööturgudel. Töötuse ja tööhõivemäärade ruumilist sõltuvust uuritakse kasutades andmeid 306 NUTS2 piirkonna kohta Euroopas. Ruumilise autokorrelatsiooni hinnangud näitavad, et piirkondlikud tööturud Euroopas on ruumiliselt klasterdunud, st kõrge (madala) töötuse/hõivemääraga piirkonnad on ümbritsetud piirkondadega, kus on kõrge (madal) töötus/hõivemäär.

Ruumilist sõltuvust uuritakse ka erinevate ökonomeetriliste mudelite abil, mis arvestavad sõltuva muutuja, sõltumatute muutujate ja vealiikme kaudu toimivaid ruumilisi mõjusid. Selleks, et teha kindlaks, kas ruumiline seos on seletatav sarnaste omadustega piirkondade klasterdumisega, kontrollitakse mudelites erinevate taustatunnuste suhtes. Piirkondlikke tööturge mõjutavad tegurid valitakse põhinedes töötuse tasakaalu ja tasakaalutuse teooriatele (ingl keeles *equilibrium and disequlibrium view*, Marston 1985).

Analüüsi tulemused näitavad, et varasemalt täheldatud töötuse ja tööhõivemäärade regionaalsete erinevuste põhjuseks on vanuselise struktuuri, majandusstruktuuri, inimkapitali ja riigi tasandi institutsioonide erinevused piirkondades. Kuid ka nende tegurite suhtes kontrollides on töötuse ja tööhõive määrade ruumiline sõltuvus statistiliselt oluline. Ruumiline sõltuvus on positiivne nii töötuse kui ka tööhõive määra osas, mis näitab, et piirkondade vahelised koostöö efektid domineerivad mõjusid, mis tulenevad konkurentsist tööjõu pärast erinevate piirkondade vahel. Tulemused näitavad, et ühe regiooni tööhõive/töötuse määr on otseselt mõjutatud teistes piirkondades toimunud tööhõive/töötuse määra muutustest, aga ka teistes piirkondades ilmnenud šokkidest. Seega esineb piirkondlikel tööturgudel olulisi ruumilisi ülekandeid. Uuringu tulemusel ei leitud empiirilisi tõendeid selle kohta, et ruumiline mõju leviks demograafiliste, nt noorte osakaalu muutuste kaudu, muutuste kaudu majandusstruktuuris või muutuste kaudu inimkapitalis. Ruumilise sõltuvuse täpne mehhanism jääb huvitavaks küsimuseks, mida uurida tulevases uurimistöös

Uuringu tulemused annavad olulist informatsiooni, mida arvestada tööturu ja regionaalpoliitika meetmete väljatöötamisel Euroopas. Regionaalse töötuse vähendamiseks ja tööhõive suurendamiseks kavandatavate poliitikameetmete puhul tuleks arvestada kohalike tööturgude vaheliste ruumiliste seostega ning seetõttu peaksid naaberregioonid meetmete arendamist omavahel kooskõlastama. Näiteks infrastruktuuri ja transpordiühenduste arendamine naaberregioonide vahel võib olla kasulik mõlemale piirkonnale, kuna see võimaldab töötajate ja töökohtade efektiivsemat sobitamist. Kokkuvõttes nõuab kohalikel tööturgudel valitsevate erinevuste vähendamine tihedat koostööd erinevate regioonide vahel.