

UNIVERSITY OF TARTU  
Faculty of Social Sciences  
School of Economics and Business Administration

Laura Helena Kivi

**SPATIAL DEPENDENCE IN EUROPEAN LABOUR MARKETS**

Master's thesis

Supervisor: Professor Tiiu Paas

Tartu 2017

Name and signature of supervisor.....

Allowed for defence on.....

(date)

I have written this master's thesis independently. All viewpoints of other authors, literary sources and data from elsewhere used for writing this paper have been referenced.

.....

(signature of author)

## **Acknowledgements**

I am very grateful to my supervisor professor Tiiu Paas for her support and guidance throughout the whole writing process. I would like to thank Jaan Masso and Kadri Ukrainski for useful comments and thoughts on the earlier versions of this thesis and Sven-Kristjan Bormann for a thorough feedback on the thesis. Finally, I am sincerely thankful to my family and friends for their support during the writing process.

## **Abstract**

The study investigates spatial dependence of unemployment and employment rates in the European labour markets relying on the Eurostat NUTS2 level data of 306 European regions. Spatial dependence is explored by using spatial error, spatial lag and spatial autoregressive model with spatial autoregressive disturbances. The results of the study confirm the importance of spatial interaction in regards to regional labour markets in Europe. The findings show that regional labour markets in Europe cluster in space, i.e. 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 the regional labour markets exist. (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 population of youth, differences in industrial structure or difference in human capital.

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

JEL classification: C21, E24, R23

## **1. Introduction**

Regional disparities in labour markets development between different European countries have been long noted. While institutional differences between different countries might explain the disparities in unemployment and employment rates at a country level, there still exist large differences in given rates among regions within the same country. Based on Eurostat database, the regional unemployment rates have been remarkable varied during the recent 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 in 1996–2007, a polarization has followed in period 2007–2013. The great disparities in unemployment and employment result in differences in the income of individuals in the region and therefore

lead to higher inequality between regions. Furthermore, as Taylor (1996) states the reduction of regional unemployment differences 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 regional variability of unemployment by differences in various factors, such as demographic factors, human capital, amenities, industrial composition, unemployment benefits (see for an overview Elhorst, 2003). However labour market participants are not restricted to work only in their resident region. While looking for employment opportunities, workers consider also neighboring labour markets. Overman and Puga (2002) state that regional unemployment is related much more to the neighboring regions than to the other regions within the same country. Positive spatial dependence in regional unemployment rates has been noted for different countries (some examples are 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. Thus previous empirical evidence indicates, that while investigating regional labour market differentials, it is essential to investigate the role of spatial dependence.

While spatial dependence has been included in the analysis in case of unemployment rates in different studies, employment rate spatial dependence has been investigated in a very few cases. Overall the spatial dependence in regional unemployment is found to be positive while there are somewhat mixed results in terms of the sign of the relationship of employment rate (see e.g. Pavlyuk 2011, Lewis *et al.* 2011, Mayor and López 2008). The positive spatial dependence in unemployment rates is mostly explained in the literature by the commuting and migration of the residents between different regions (see e.g. Molho 1995, Patacchini and Zenou 2007). On the contrary, in case of employment, competition among regions for qualified workers can result in negative spatial dependence (see e.g. Mayor and López 2008). However, agglomeration and cooperation of industries in different regions can also lead to positive spatial dependence in terms of employment (see e.g. Lewis

*et al.* 2011). Therefore, there is a clear gap in the literature in terms of employment rate spatial dependence, as it is not certain, whether cooperation effects, which lead to positive dependence, or competition effects, which lead to negative dependence, dominate on the regional labor market.

The aim of this study is to investigate spatial dependence at the regional level in Europe both in terms of unemployment and employment rates. No previous studies have been conducted where both unemployment and employment rate spatial dependence has been analyzed with spatial models for the same dataset. It is important to investigate both unemployment and employment rate spatial dependence to better understand the processes of regional labor markets. While both unemployment and employment are related to the overall equilibrium in the labour market, their spatial dependence is likely to differ in terms of strength of the relationship and possibly also sign. In the case, where commuting and migration lead to positive spatial dependence in unemployment rates, but competition for labour force 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 from this study provide additional information for developing new labour market policy measures in Europe taking into account possible spatial relations between labour markets of countries and regions.

This study uses European NUTS 2 regions data on unemployment and employment rates. The analysis is based on the results from spatial regression models. Specifically, spatial lag model, spatial error model and spatial autoregressive model with spatial autoregressive disturbances 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, i.e. regions with high (low) (un)employment rate are surrounded by regions with high (low) (un)employment rate. The results confirm 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 the regional labour markets. Interestingly, spatial dependence between regional labour markets in Europe has been fairly stable during recent decade. No evidence was found that the spatial effects work through the differences in demographics, such as the share of population of youth, differences in industrial structure or differences in human capital.

The paper is structured as follows. Section 2 presents the literature review. Section 3 explains the used 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 of investigating labour market differentials. Equilibrium and disequilibrium view (Marston 1985) that explain the disparities of the regional labour markets are examined. Factors that affect the adjustment of the regional labour market according to those views are analyzed. While the factors affecting unemployment rate are noted based on theoretical considerations, the results from empirical studies concerning the impact of those factors are also included.<sup>1</sup> In the second subsection attention is turned to specifically 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, first part of the second subsection is dealing with the spatial dependence in unemployment rates and the mechanism behind, while the second part is focusing on the employment rate spatial dependence.

---

<sup>1</sup> 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.

## **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: 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 factors to their analysis to determine whether the 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 high unemployment rate in a region is compensated by some other factors (higher wages, amenities, lower overall costs of living, industry composition). In this view external shocks affect unemployment rates only for a short period of time, allowing it to converge back into its underlying mean value (Semerikova 2015). Marston (1985) claims that if the unemployment is of equilibrium nature, then government's attempts to reduce regional disparities are useless as it is impossible to reduce the regional unemployment rate in the long term.

In equilibrium view most of the factors affecting 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 unemployment rate as it decreases labour demand and increases labour supply. In 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 average wage. On the contrary, Badinger and Url (2002) report negative relationship, which is explained by the fact that opportunity costs to stay unemployed are higher in an area with higher average wage.

Unemployment disparities may also originate from differences in amenities. As Aragon *et al.* (2003) state, according to equilibrium view, areas with more pleasant climate, active cultural life or better infrastructure are expected to exhibit higher unemployment. Lower housing costs are also sometimes seen as compensation factor for high unemployment (Semerikova 2015). Marston (1985) confirms for the US data that amenities such as abundance of parks per capita, pure air to breathe, and mild climate are associated with a higher rate of unemployment in the area. Population density is sometimes included as a measure of the quality of life. Areas with low population density can be seen as more favorable living environments as they tend to have stronger social networks (Badinger and Url 2002). Aragon *et al.* (2003) however report positive relationship for France 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 disequilibrium view below.

Different industrial composition 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 not significant (e.g. Semerikova 2015, López-Bazo *et al.* 2002, López-Hernández 2013, Filiztekin 2009, Niebuhr 2003, Diaz 2016). Industry diversity indexes have also been used as measure for industrial composition. The greater is the diversity of industry, the more easily can employment reductions in one sector be absorbed by the other sectors (Diaz 2016). However empirical results in this aspect have not been found significant (e.g. Mitchell and Bill 2004, Diaz 2016).

In disequilibrium view regional unemployment rates should become equal between the regions in the long run (Aragon *et al.* 2003). Workers from high unemployment areas would migrate to other regions and firms would relocate to the high unemployment areas with a perspective of a possible workforce, leveling out the regional differences. However

speed of adjustment tends to be slowed down by the restrictions on mobility on both sides: workers experience the costs 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). Thus labour markets do not manage to reach this equal unemployment rate before a new shock (e.g. a factory closure) hits the labour markets (Diaz 2016). Contrary to the equilibrium view, where long run differentials could not be reduced by the governmental policies, here introducing more flexibility to labour markets and reducing migration costs could help to increase the speed of adjustment, hence reduce disparities in regional unemployment rates in long term.

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

Age structure of the population is thought to 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 generation (Aragon *et al.* 2003). Filiztekin (2009) and Diaz (2016) confirm, based on Turkish and Colombian regional data respectively, that the share of young people in working age population is negatively related to unemployment. However some studies also find that regions with 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 share of 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. Firstly, 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 more demanded 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 for German data as both share of persons without professional education and the ones with university education have 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 negative effect dominates, one would expect higher share of high skilled individuals to increase the speed of adjustment and lower regional unemployment rates.

Employment growth reduces unemployment almost 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 of the 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 equilibrium view, it is likely to affect the speed of adjustment and should be therefore considered also under 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 effect of unemployment benefits is straightforward: they slow down the adjustment process by increasing the reservation wage of the unemployed and decreasing the motivation to find a new job in home region or to migrate for that purpose. The results of Badinger and Url (2002) and Marston (1985) are in line with the aspect.

The mobility of the workers and therefore the speed of adjustment of the labour market is restricted by the magnitude of migration costs. Most important in this aspect is the structure of the housing market: e.g. housing prices, share of apartments owned/rented out. Badinger and Url (2002) find that regions with 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 to find another housing, with subsidized cheaper prizes, is low.

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

## **2.2. Spatial dependence in regional labour markets**

While above mentioned 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 for Germany (Semerikova 2015), Japan (Kondo 2015), UK (Molho 1995, Patacchini 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 detected spatial autocorrelation was positive, meaning that the neighbors 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

neighboring regions have similar local characteristics, e.g. in terms of skill composition of the population or industrial structure. However almost all of the named studies (except Kondo 2015, Pattacchini and Zenou 2007) also analyzed spatial regression models, adding different controls to account for the various local characteristics, and still found significant spatial effects.

Although mentioned studies 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 is commuting and migration across neighboring regions as people look for work both in area they live in and in the areas they do not. Important contribution to the studies of this aspect have been made based on UK data by Molho (1995), who analyzes 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 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 to neighboring areas that increases in time. The fact that spillover effect is stronger after a time lag points to migration behavior: when higher labour demand is noted in the neighboring labour market, workers need time to make arrangements (e.g. find appropriate housing, school for their children etc.) for relocation. Molho (1995) also identifies highly localized effects that point to commuting. In line with those results is the study of 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 behavior 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 a 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 to spatial dependence in manufacturing sector in South Carolina

counties. Change in manufacturing employment is found to have positive relationship with neighboring counties' employment change. Here the sign of the relationship is likely to results 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 spatial dependence of employment change to be slightly negative. The results of the spatial dependence of employment are therefore mixed in terms of sign of dependence. The mixed results might be explained by differences in data, e.g. the size differences of the geographic units used in different studies, focusing only on one industrial sector employment (e.g. manufacturing sector employment), or they might indicate different forms of spatial interaction, e.g. competition among regions for qualified workers resulting in negative effect versus agglomeration and cooperation of industries in different regions resulting in positive effect.

In summary, previous studies have found positive spatial dependence in regional unemployment rates and somewhat mixed result on the sign of dependence in case of employment change and rates. There is a clear gap in literature in terms of employment rate spatial dependence, as one cannot be certain whether regions with high employment rates really affect the neighboring regions employment rate negatively or the negative relationship shown in the above mentioned studies is just an empirical irregularity. Also although there are rather many studies, that analyze spatial dependence at regional level for one country, the most recent study using spatial regression models to analyze regional unemployment spatial dependence for many neighboring 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 workers in border regions are in the light of European Union free labour movement principle also likely to seek work opportunities in neighboring regions across the national border.

This study is focusing on empirical analysis of spatial dependence in regional labour markets both for employment and unemployment rates and for the regions in European countries, by involving the new data sets that have recently become available. A spatial

econometrics modeling approach is applied here in order to account for the inter-regional differences.

### **3. Method and Data**

#### **3.1. Data**

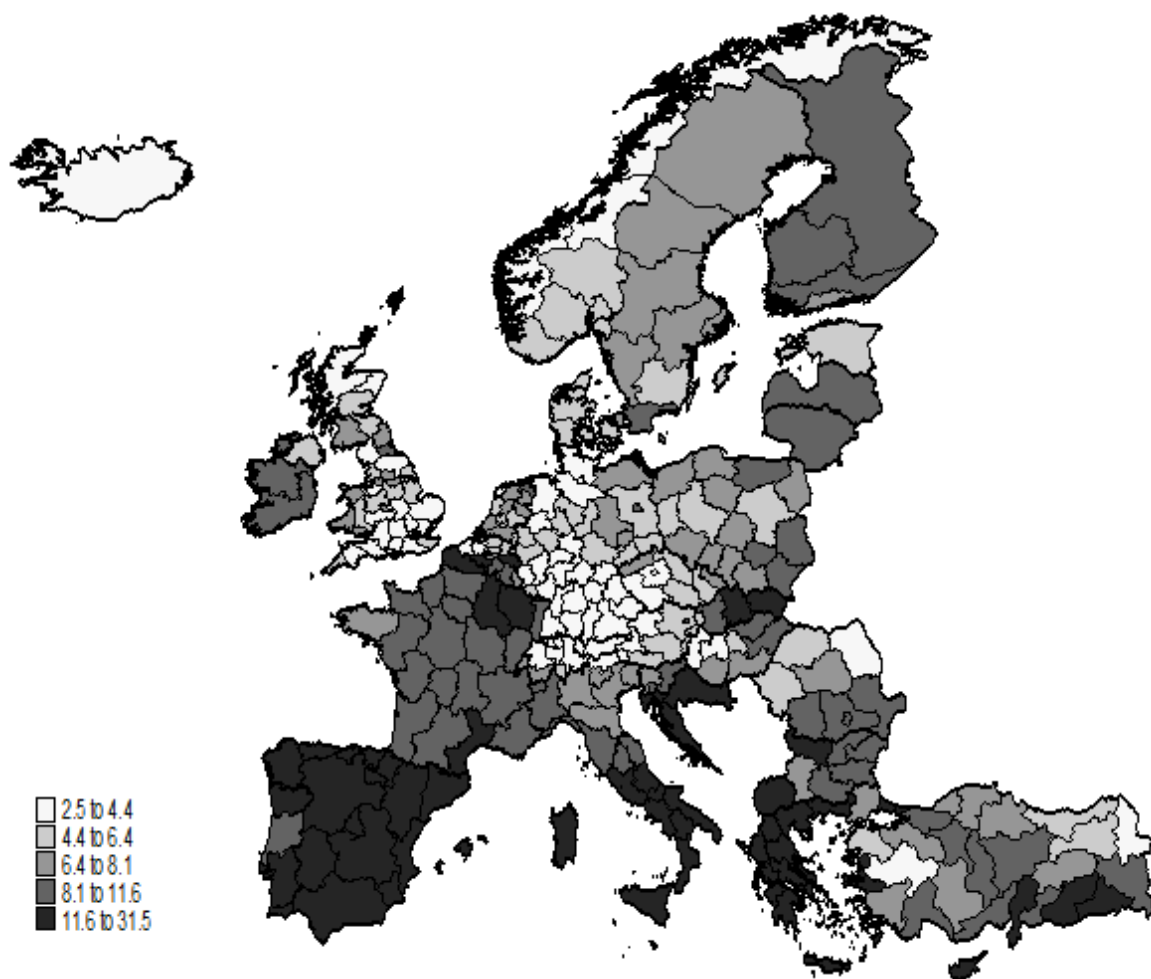
The data used in current study is provided by Eurostat database. Regional unemployment rate is defined as number of unemployed persons to the number of persons in the economically active population (i.e. sum of employed and unemployed). Unemployed persons comprise persons aged 15-74 who were: 1) without work during the reference week; 2) currently available for work; 3) actively seeking work or who had found a job to start within a period of 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 indicator is based on the EU Labour Force Survey and is in accordance with the International Labour Organization (ILO) definition of unemployment. Regional employment rate is defined as number of employed persons to the population of the age group 15-64 (i.e. working age population). This indicator is also based on the EU Labour Force Survey.<sup>2</sup>

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. NUTS 2 level is defined as basic regions for the application of regional policies (NUTS overview). NUTS 2013 classification is used in the current study. The study uses cross section data for 306 regions in Europe for the year 2015, the newest data available.<sup>3</sup>

---

<sup>2</sup> Description of the variables and source of exact datasets used are given in appendix 1.

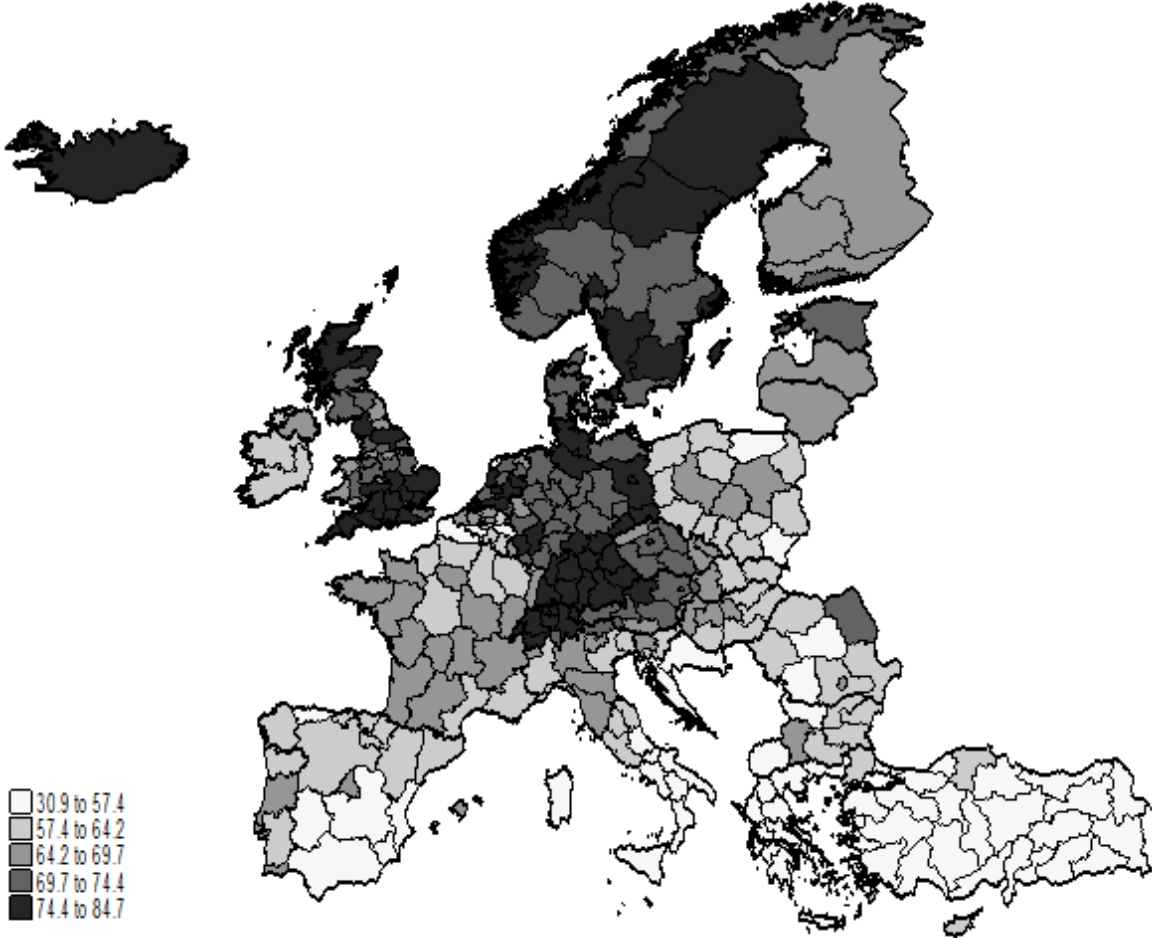
<sup>3</sup> List of countries and number of regions included in the analysis is shown in appendix 2.



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

Figure 1 presents the unemployment rates in European NUTS 2 regions. It can be noticed that regions with similar unemployment rates are rather concentrated. Although most of the clustering seems to be inside of country borders (e.g. high unemployment in south of Italy in comparison with middle and north of Italy), also some cross border similarities can be seen. For example border regions in south of Germany, Switzerland, Austria and Czech Republic have similar low unemployment rates. Whether this is a sign of cross-border interaction between the regional labour markets or is 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.



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

Employment rates in NUTS 2 regions in Europe are displayed on figure 2. Similar to unemployment, clustering can be seen inside country borders (e.g. Spain, France), but also across the national borders. In accordance with the results on unemployment rates, regions in south of Germany have similar employment rate values with their neighbors across the border in Switzerland and Austria. Again it must be noted that similarities in neighboring regions seen in raw data can be partially a sign of interaction across regions and partially be

accounted for cross-country differences or regional characteristics (e.g. demographics, industrial structure).

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: share of population aged 25-64, who have obtained 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, 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 whole 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 the regions in different European countries, country dummies are added as control variables.

### 3.2 Spatial autocorrelation

To account for spatial dependence spatial weight matrix  $W$  can be used. Spatial weight matrix determines the structure and the intensity of spatial dependence between the regions. There are various ways to specify 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.

A frequently applied and simplest weight specification is a binary spatial weight matrix such that the elements of the matrix  $w_{ij} = 1$  if regions  $i$  and  $j$  share a border and  $w_{ij} = 0$  otherwise (e.g. Diaz 2016). However simple binary matrix is not always appropriate as it assumes that spatial autocorrelation only occurs between nearest neighboring 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 distances between the spatial units. One advantage of this type of matrices is, that they allow all the weights between the regions to be positive and thus do 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 account for the spatial interaction also between the regions that are not direct neighbors, distance-based matrix is used. The elements of the matrix used in the current analysis are constructed as the inverse values of distances between the geographic centers of the regions:

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

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, i.e. the effect of weighting operation can be interpreted as averaging over neighboring values (Elhorst 2014).

To determine whether spatial clusters of low and high unemployment and employment exist, measures of spatial autocorrelation are used. If similar parameter values, whether high or low, are spatially clustered, then positive spatial autocorrelation can be detected in data. Spatial closeness of dissimilar values is indicated by negative spatial autocorrelation. 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 variable of interest in regions  $i$  and  $j$ ,  $\bar{x}$  is the average of  $x$  over the  $N$  regions and  $w_{ij}$  is the element of a spatial weights matrix  $W$  summarizing the interaction between regions  $i$  and  $j$ .

Moran's I takes values in the interval of  $[-1;1]$ . Positive and significant Moran's I indicates spatial clustering of similar parameter values.

The estimated Moran's I for unemployment and employment rate are shown in table 1. There is a significant positive spatial correlation both for unemployment and employment rates. Thus there is evidence of spatial clustering and potential spatial interaction. It should be noted that spatial autocorrelation is stronger and also positive in the case of employment rates, which is interesting in the light of somewhat mixed results from earlier literature on the sign of dependence. The positive sign of spatial autocorrelation for 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.

**Table 1.** Moran's spatial autocorrelation index for unemployment and employment rate

	Moran's I	z
Unemployment rate	0.18***	33.444
Employment rate	0.262***	48.175

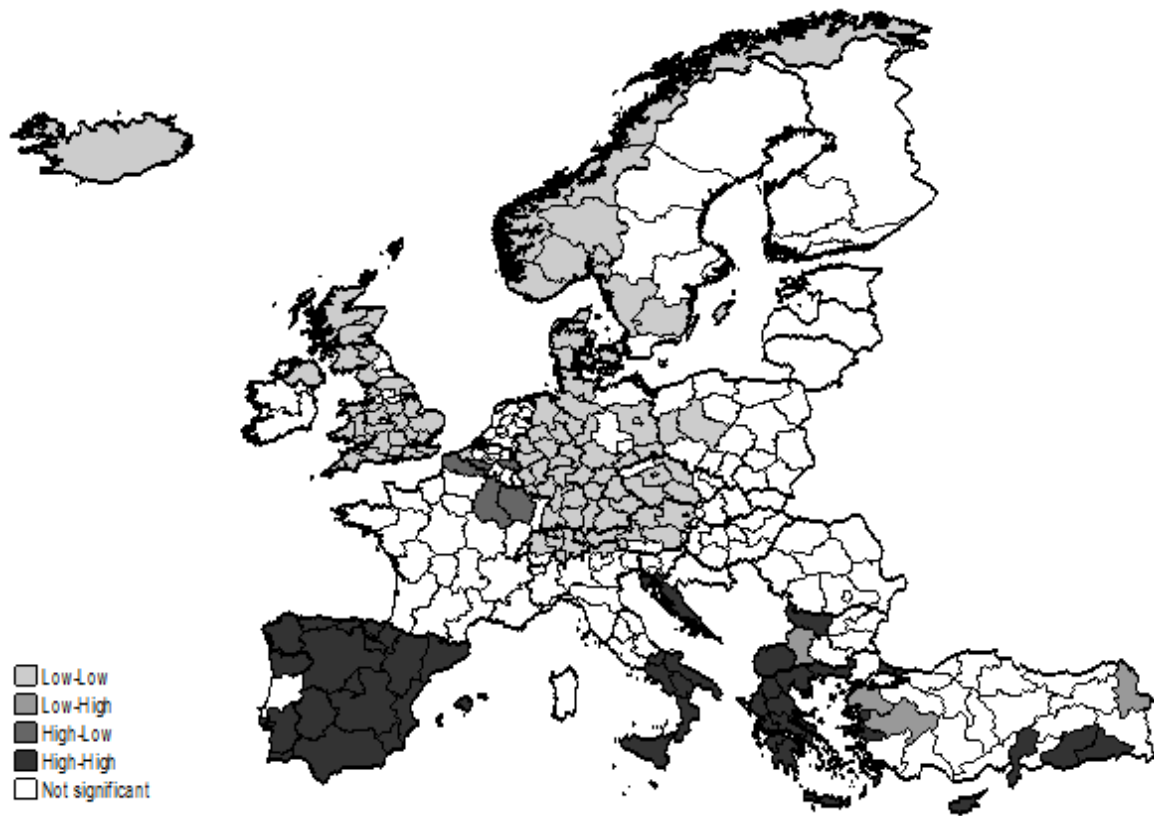
Note: \*\*\* indicates the values are significant at 1% level. N=306

The presented Moran's I gives the value of 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 neighbors. To measure the local spatial autocorrelation, local indicators of spatial association (LISAs) for Moran's I (Anselin 1995) are defined as follows:

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

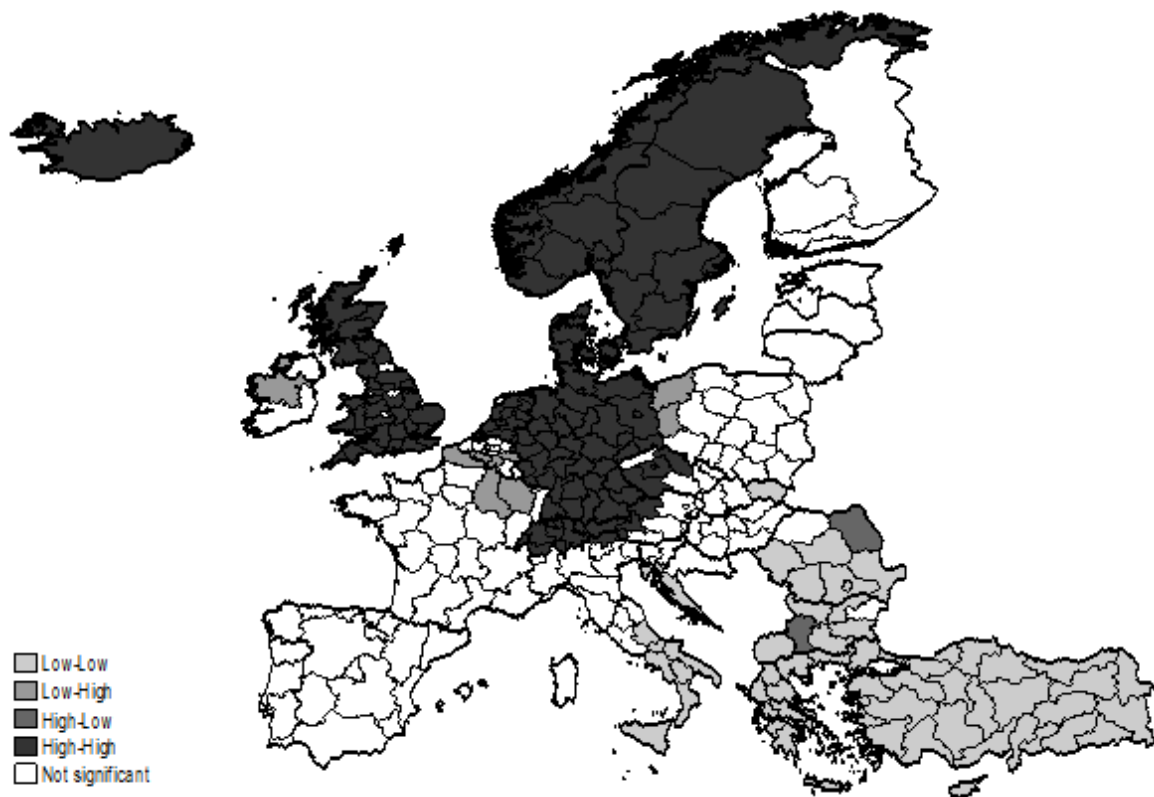
where  $N$  is the number of regions indexed by  $i$  and  $j$ ,  $x_i$  and  $x_j$  are the variable of interest in regions  $i$  and  $j$ ,  $\bar{x}$  is the average of  $x$  over the  $N$  regions and  $w_{ij}$  is the element of a spatial weights matrix  $W$  summarizing the interaction between regions  $i$  and  $j$ .

Figure 3 displays the areas with significant LISAs for unemployment rates. The darkest and lightest 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 neighbors 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 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 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 neighbors. This association is likely to at least in part be the result of commuting and migrating across national borders, especially between the areas that share a common language.



**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 on figure 4. Again clusters inside country borders are visible, e.g. in UK, Norway and Sweden. Although spatial association in southern Italy and Greece were also found significant in terms of unemployment rates, here a bigger cluster is forming with regions in Greece, Turkey, Bulgaria and Romania all exhibiting positive spatial autocorrelation with their neighbors. It should be noted that regions in Spain and Portugal are not found significant in terms of spatial association for employment rates. Similar to figure 3, regions in Germany and its neighbors stand out as the ones with significant spatial association. In this case also regions in Netherlands are part of the cluster.



**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 neighbors, though positive relationship clearly dominates. While some regions form clusters inside countries, data on regions in Germany, Switzerland, Austria, Denmark, Poland (for unemployment rates) and Netherlands (for employment rates) point to the existence spatial association across national borders. To determine whether the spatial association can be explained by similar regional characteristics or is partly a sign of cross border interaction (e.g. in form of commuting), 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). These three types of interaction effects can explain why observation in one region may be dependent on observations at other regions.

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 neighboring units. Endogenous interaction effects reflect the substantive form of spatial autocorrelation. In case of exogenous interaction effects dependent variable of a particular unit depends on 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 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 nuisance form of spatial dependence.

The most commonly used spatial models are spatial lag model (SLM), also known as spatial autoregressive model (SAR), that accounts for the endogenous interaction effect and spatial error model (SEM), that accounts for the interaction effect in error terms, both presented in the seminal book by Anselin (1988). Ignoring spatially lagged dependent variable may lead to biased and inefficient estimates. Ignoring spatially correlated errors may result in inefficient estimates.

Spatial error model (SEM) takes the form:

$$\mathbf{y} = \alpha \mathbf{1}_N + \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

$$\mathbf{u} = \lambda \mathbf{W} \mathbf{u} + \varepsilon$$

where  $\mathbf{y}$  is a  $(N \times 1)$  vector of dependent variables,  $\mathbf{X}$  is a  $(N \times k)$  matrix of  $k$  explanatory variables,  $\mathbf{1}_N$  is a vector of ones,  $\lambda$  is the spatial autocorrelation coefficient,  $\mathbf{W}$  is the  $(N \times N)$  spatial weight matrix and random term  $\varepsilon \sim N(0, \sigma^2)$ .

Spatial lag model (SLM) is defined as:

$$\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \alpha \mathbf{1}_N + \mathbf{X} \boldsymbol{\beta} + \varepsilon$$

where  $\mathbf{y}$  is a  $(N \times 1)$  vector of dependent variables,  $\mathbf{X}$  is a  $(N \times k)$  matrix of  $k$  explanatory variables,  $\mathbf{1}_N$  is a vector of ones,  $\rho$  is the spatial autoregressive coefficient,  $\mathbf{W}$  is the  $(N \times N)$  spatial weight matrix and error term  $\varepsilon \sim N(0, \sigma^2)$ .

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

$$\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \alpha \mathbf{1}_N + \mathbf{X} \boldsymbol{\beta} + \mathbf{u}$$

$$\mathbf{u} = \lambda \mathbf{W} \mathbf{u} + \varepsilon$$

where the notion is as described above.

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

$$\mathbf{y} = \rho \mathbf{W} \mathbf{y} + \alpha \mathbf{1}_N + \mathbf{X} \boldsymbol{\beta} + \mathbf{W} \mathbf{X} \boldsymbol{\theta} + \varepsilon$$

where  $\mathbf{y}$  is a  $(N \times 1)$  vector of dependent variables,  $\mathbf{X}$  is a  $(N \times k)$  matrix of  $k$  explanatory variables,  $\mathbf{1}_N$  is a vector of ones,  $\rho$  is the spatial autoregressive coefficient,  $\mathbf{W}$  is the  $(N \times N)$  spatial weight matrix,  $\boldsymbol{\theta}$  is a  $(N \times 1)$  vector of coefficient estimates of spatially lagged values of explanatory variables and error term  $\varepsilon \sim N(0, \sigma^2)$ .

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

### 3.4 Direct and indirect effects

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

It is possible to rewrite the spatial lag model as:

$$(I - \rho W)y = X\beta + \iota_N\alpha + \varepsilon$$

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

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:

$$\frac{\partial y_i}{\partial x_{jr}} = (I - \rho W)^{-1}\iota_N\beta_r = S_r(W)_{ij}$$

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 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 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 2 reports the estimates from OLS, SEM, SLM and SARAR models explaining the regional unemployment rate differentials.<sup>4</sup> 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. Anselin *et al.* (1996) present Lagrange multiplier (LM) tests to for both types of interaction. According to the decision rule spatial dependence is of the spatial lag form if 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. Opposite indicates that the spatial dependence is of the spatial error form. The results from LM-tests are reported in table 2 and reject the null of no spatial dependence against the both forms of spatial dependence. The results from the robust LM-test indicate that SLM model would be the best specification compared to SEM model. Based on Akaike information criterion (AIC) SARAR model seems to be the best choice, however Bayesian information criterion (BIC) is marginally lower for SLM model.

The estimates show interesting features, most of them consistent with the theoretical background and empirical literature that analyzes unemployment rates in different regions and countries, presented in previous section. It should be noted that the coefficient estimates differ only marginally between the different models. All of the estimates are also

---

<sup>4</sup> Results for 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 SDM model does not provide additional information compared to SLM model. Thus SDM results are not commented further.

statistically significant, except the share of manufacturing. A higher share of young people is related to 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 younger generation has in terms of entering the labour market.

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

<b>Unemployment</b>	<b>OLS</b>	<b>SEM</b>	<b>SLM</b>	<b>SARAR</b>
Youth	0.329** [0.074]	0.356** [0.071]	0.340** [0.066]	0.363** [0.068]
Services	0.155** [0.033]	0.158** [0.030]	0.165** [0.029]	0.168** [0.029]
Manufacturing	-0.047 [0.036]	-0.036 [0.034]	-0.033 [0.032]	-0.023 [0.032]
Higher education	-0.165** [0.034]	-0.172** [0.032]	-0.167** [0.030]	-0.176** [0.031]
constant	-0.059 [0.033]	-0.061 [0.033]	-0.135** [0.030]	-0.139** [0.030]
$\lambda$		0.883** [0.120]		0.813** [0.173]
$\rho$			0.929** [0.069]	0.918** [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 <sup>2</sup>	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 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 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 presented in the last section 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 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 investments. Therefore in short term services sector is more flexible with the movements of unemployment than the manufacturing sector. Another explanation for the results can be that services and manufacturing sector are internally heterogeneous, i.e. manufacturing and services sector include wide range of different skill level jobs. This opportunity will be investigated below.

The share of population with higher education has an expected relationship with the unemployment rate. In regions with higher share of high skilled people, the unemployment rate is lower. This is in line with the earlier empirical results (e.g. Diaz 2016, López-Bazo *et al.* 2002, López-Hernández 2013) and also with theory that points to the faster speed of adjustment of highly educated people. Most of the coefficients of country dummies are found to be significant in included models, 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 part of the result are the estimates of spatial autocorrelation and spatial autoregressive coefficients that give an answer to the earlier question about the existence and statistical significance of spatial interactions between the regional labour markets. While the preliminary analysis carried out in the previous section did point to significant spatial associations among the regional labour markets in Europe, it was not clear if the significant relationship was only reflecting the spatial clustering of regions with similar characteristics or pointing to the effects of spatial interaction, e.g. in form of commuting. The presented models control for age structure, human capital, industrial

structure and the role of institutions and still find the spatial autocorrelation coefficient  $\lambda$  in SEM model and the spatial autoregressive coefficient  $\rho$  in SLM model to be significant. SARAR model spatial coefficient estimates are also both significant and similar to the respective coefficient estimates in SEM and SLM model. It is worth to note that the significance of both spatial coefficients shows that both substantive and nuisance form 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 exists spillovers across the regional labour markets, which can express itself, for example, in a form of workers commuting from one region to another.

The estimates from different models for employment rate are presented in table 3. Similarly to unemployment rate the LM-tests reject the null of no spatial dependence in OLS model residuals. The LM-tests indicate that SLM would be the best model compared to SEM model. Based on the information criteria the most general SARAR model is the preferred specification. As in case of unemployment rates, the coefficient estimates are similar across the models. All of the estimates are also statistically significant, except the share of manufacturing, which was also the case with unemployment rates. The areas with higher share of young population have on average lower employment rates. The relationship is stronger here than the positive relationship in terms of unemployment rates. This results is expected as employment rate is defined as the share of employed in whole working age population, while unemployment rates looks at share of unemployed in labour force. Therefore an area with a high share of young people also tends to have a higher share of inactive population as young people are often engaged with studies rather than working. The results in terms of industrial composition are in line with the results in case of unemployment rates. Manufacturing share is also not found significant here and the share of services has a negative relationship with regional employment rates. The positive relationship of higher share of higher educated is as expected and reflect the fact that more educated people are more efficient in finding jobs and also more demanded in the labour market. Country dummies were included in all of the models to account for cross country

differences in institutions and legislation. 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 the regional labour market differences.

**Table 3.** OLS, SEM, SLM and SARAR estimates of regional employment rate determinants

<b>Employment</b>	<b>OLS</b>	<b>SEM</b>	<b>SLM</b>	<b>SARAR</b>
Youth	-0.733** [0.108]	-0.753** [0.103]	-0.689** [0.095]	-0.708** [0.097]
Services	-0.244** [0.048]	-0.249** [0.044]	-0.259** [0.042]	-0.263** [0.042]
Manufacturing	0.045 [0.053]	0.024 [0.049]	0.005 [0.047]	-0.012 [0.046]
Higher education	0.314** [0.050]	0.330** [0.046]	0.312** [0.043]	0.332** [0.044]
constant	0.926** [0.048]	0.925** [0.052]	0.279** [0.058]	0.289** [0.063]
$\lambda$		0.921** [0.080]		0.891** [0.083]
$\rho$			0.941** [0.058]	0.933** [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 <sup>2</sup>	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 show spatial interactions also in case of employment rate. The spatial autocorrelation coefficient in SEM model and the spatial autoregressive coefficient in SLM model are both significant and of high value. It should be noted that also in SARAR model both estimates of spatial coefficient are significant and their values are respectively similar to those in case of SEM and SLM model. Therefore, both substantive and nuisance form of

spatial autocorrelation exists for employment rates. While based on earlier studies (e.g. Pavlyuk 2011, Lewis *et al.* 2011, Mayor and López 2008) positive spatial dependence could not be confirmed in terms of employment rates, these results give support to the existence of positive spillovers also in 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 4. The direct effects estimates are similar to the coefficient estimates of SLM model (see table 2). As before, the coefficient of the share of manufacturing is not found significant. The direct effects estimates for other variables lead also to the same conclusions as the coefficient estimates for SLM model.

The average indirect effect shows the effect changes in each explanatory variable in one region have on unemployment in other regions. It is important to note, that this measure shows the cumulative impact over space. Therefore indirect effect is often estimated to be higher than the direct effect. Interestingly, indirect effects are not found to be significant in this case. Thus 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 4.** Direct and indirect effect estimates for unemployment rate (based on SLM)

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

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

As explained above, earlier studies have found that one of the main mechanisms causing spatial dependence are commuting and migration across neighboring regions as people look for work both in area they live in and in the areas they do not (see Molho 1995, Pattacchini and Zenou 2007). Investigating commuting and migration as causes for spatial dependence is however limited in current analysis by data unavailability for NUTS 2 level. While Eurostat exhibits dataset on commuting on NUTS 2 level, data for many regions is missing. For migration behavior, data on net migration and crude rates of net migration is available. Net migration is the difference between immigration and emigration and can be of small value, when both immigration and emigration are both of high similar value. Therefore net migration does not capture the intensity of migration and is thus unsuitable for the current analysis. Consequently, the exact mechanism of spatial dependence is left for future research in the hope of improved data availability. Commuting and migration will be used as illustrative examples of spatial interactions based on earlier studies.

The direct and indirect estimates of employment rates are presented in table 5. Direct effects are again slightly higher than the coefficient estimates of SLM model (see table 3). The indirect effects are not found statistically significant for employment rates. Therefore also for employment rates spatial effects do not seem to work 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 5.** Direct and indirect effect estimates for employment rate (based on SLM)

Employment	Direct effect	Indirect effect
Youth	-0.756** [-6.39]	14.413 [-1.13]
Services	-0.282** [-5.60]	-5.440 [-1.09]
Manufacturing	0.006 [0.12]	-1.095 [0.08]
Higher education	0.342** [6.21]	6.606 [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 European Union in 2004, SEM, SLM and SARAR models have been estimated for the years 2004, 2008, 2011 and 2015. For 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 6. Firstly, spatial dependence in all the given years is rather similar. By comparing the years 2004 and 2015, it can be seen that spatial dependence is slightly higher. However, the years in between exhibit a slightly lower dependence. With the presence of spatial lag dependence, spatial error dependence turns out to be significant only for the most recent year.

**Table 6.** Spatial coefficients for years 2004, 2008, 2011 and 2015 for unemployment rate

Unemployment		2004	2008	2011	2015
SEM	$\lambda$	0.803**	0.794**	0.742**	0.856**
SLM	$\rho$	0.907**	0.863**	0.851**	0.926**
SARAR	$\lambda$	0.64	0.654	0.369	0.720**
	$\rho$	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 7). 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.

These results may seem somewhat puzzling at first, as one might expect the spatial dependence, as a sign of European labour market's integration, to grow considerably over the years. However there are also factors that lessen the need to commute or migrate for work into neighboring regions, and therefore presumably prevent the spatial dependence to grow. One of the reasons could be the enhanced possibilities for remote working. The modern communication possibilities enable working from distance without having the need to migrate. Another possible explanation is that difference in average wages and overall living standards between the different European regions have decreased compared to the

time of eastern enlargement of EU. Therefore, the gain from migrating or commuting for work has also diminished.

**Table 7.** Spatial coefficients for years 2004, 2008, 2011 and 2015 for employment rate

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

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

## 5. Robustness checks

To check for the robustness of results in time, panel data models were used. Data for years 2004, 2008 and 2015 for 253 NUTS 2 regions were used.<sup>5</sup> 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 to investigate the impact of time invariant explanatory variables. SARAR model is excluded, while the random effects variant of the SARAR model can be written as a special case of SLM specification (Belotti *et al.* 2016).

The results for panel data models on unemployment rate are presented in appendix 4. All of the coefficient estimates, also the share of manufacturing, are significant in these models. While the exact values of coefficients are somewhat different than for earlier results, the signs of coefficients are in accordance with the earlier results (see table 2). Overall the coefficient estimates lead us to the same conclusions as the results on cross-sectional data. Particularly, the spatial autocorrelation and spatial autoregressive coefficient are also positive, significant and of slightly higher value as for cross-sectional data. It can be concluded that even after taking the time dimension into account, the spatial dependence, of both nuisance and substantive form, remains.

<sup>5</sup> Some of the initial 306 regions were excluded because of data unavailability for years 2004 and 2008.

Appendix 5 displays the results of panel data model on employment rates. The share of manufacturing is found significant here in pooled data model. The share of persons with higher education is not found statistically significant in SLM model. Overall the values of the coefficients are similar to the ones obtained above and lead us to similar conclusions about the relationship of changes in explanatory variables and employment rate (see table 3). The spatial autocorrelation and spatial autoregressive coefficient are also positive, significant and of slightly higher value as for cross-sectional data. Therefore the spatial dependence remains also for regional employment rates after accounting for time dimension.

While the results from the estimates for industrial composition were somewhat unexpected, the robustness of results is checked by using the subcategories of sectors. Manufacturing sector includes wide range of different skill level jobs, which has been tried to characterize by dividing the share of employment in manufacturing into two categories: high-technology manufacturing (HTM) and low-technology manufacturing (LTM). 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 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. Share of low-technology manufacturing stays insignificant. These findings again seem to point to the fact that services and here also high-technology manufacturing sector are more flexible to the changes in unemployment than low-technology manufacturing. In the case of employment rate (appendix 7), dividing manufacturing into subcategories does not have an 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 3.

## 6. Conclusions

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

The spatial dependence is explored by using different types of spatial econometrics models, that account for spatial effects working through dependent variable, independent variables and error term. The results are stable throughout the models. Set of factors is controlled for to determine whether the spatial association is explained by clustering of regions with similar characteristics. The factors affecting regional labour markets are chosen based on 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 positive both for unemployment and employment rates, indicating that cooperation effects between regions dominate competition for labour force effects. Both substantive and nuisance form of spatial dependence exists, i.e. 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 exists 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 European Union in 2004. No evidence is found that the spatial effects work through the changes in the demographics, such as the share of population of youth, changes in industrial structure or changes in human capital. The exact mechanism of spatial dependence remains and 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 to capture the interactions also between regions inside the borders of smaller countries. For example, instead of one region at NUTS 2 level, Estonia is divided to 5 regions at NUTS 3 level. Unfortunately at this date 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 the policy measures in European labour markets. 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 neighboring regions. Furthermore, shocks hitting one region will affect the unemployment and employment rates of other regions. All in all, reducing labour market differences calls for a close cooperation between the regions.

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

<b>Unemployment</b>	Pooled	SEM <sub>re</sub>	SLM <sub>re</sub>
Youth	0.495** [0.081]	0.633** [0.075]	0.454** [0.066]
Services	0.085* [0.034]	0.090** [0.031]	0.098** [0.032]
Manufacturing	-0.142** [0.038]	-0.130** [0.036]	-0.118** [0.036]
Higher education	-0.068* [0.028]	-0.135** [0.029]	-0.054* [0.025]
constant	-0.070* [0.035]	-0.059 [0.059]	-0.143** [0.033]
$\lambda$		0.978** [0.012]	
$\rho$			0.978** [0.013]
Country dummies	YES	YES	YES
AIC	-2787.93	-3078.37	-3068.59
BIC	-2639.71	-2916.25	-2906.47
R <sup>2</sup>	0.434		

Note: \* significant at 5%; \*\* significant at 1%. Standard errors are in brackets. N=759.

**Appendix 5.** Panel data models for employment rate

<b>Employment</b>	Pooled	SEM <sub>re</sub>	SLM <sub>re</sub>
Youth	-0.879** [0.092]	-0.706** [0.085]	-0.495** [0.070]
Services	-0.200** [0.038]	-0.187** [0.043]	-0.217** [0.042]
Manufacturing	0.104* [0.044]	0.05 [0.053]	0.008 [0.052]
Higher education	0.313** [0.032]	0.206** [0.041]	0.056 [0.029]
constant	0.931** [0.040]	0.922** [0.051]	0.307** [0.047]
$\lambda$		0.961** [0.021]	
$\rho$			0.955** [0.025]
Country dummies	YES	YES	YES
AIC	-2589.31	-2914.06	-2921.39
BIC	-2441.09	-2751.94	-2759.27
R <sup>2</sup>	0.7374		

Note: \* significant at 5%; \*\* significant at 1%. Standard errors are in brackets. N=759

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

<b>Unemployment</b>	<b>OLS</b>	<b>SEM</b>	<b>SLM</b>	<b>SARAR</b>
Youth	0.380** [0.073]	0.411** [0.069]	0.376** [0.066]	0.405** [0.068]
KIS	0.238** [0.051]	0.230** [0.047]	0.243** [0.046]	0.234** [0.045]
LKIS	0.146** [0.044]	0.154** [0.041]	0.155** [0.040]	0.164** [0.040]
HTM	-0.150* [0.067]	-0.135* [0.061]	-0.121* [0.060]	-0.109 [0.060]
LTM	-0.008 [0.053]	-0.006 [0.049]	-0.002 [0.048]	-0.001 [0.047]
Higher education	-0.222** [0.039]	-0.225** [0.036]	-0.222** [0.035]	-0.228** [0.035]
constant	-0.081** [0.030]	-0.082** [0.029]	-0.147** [0.028]	-0.148** [0.029]
$\lambda$		0.871** [0.131]		0.820** [0.170]
$\rho$			0.874** [0.117]	0.849** [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 <sup>2</sup>	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)

<b>Employment</b>	<b>OLS</b>	<b>SEM</b>	<b>SLM</b>	<b>SARAR</b>
Youth	-0.733** [0.112]	-0.750** [0.105]	-0.672** [0.098]	-0.692** [0.100]
KIS	-0.485** [0.078]	-0.472** [0.071]	-0.507** [0.069]	-0.492** [0.067]
LKIS	-0.101 [0.068]	-0.118 [0.062]	-0.124* [0.060]	-0.141* [0.059]
HTM	0.15 [0.102]	0.112 [0.093]	0.085 [0.090]	0.054 [0.088]
LTM	-0.012 [0.081]	-0.013 [0.074]	-0.052 [0.071]	-0.047 [0.070]
Higher education	0.457** [0.060]	0.465** [0.054]	0.458** [0.052]	0.470** [0.052]
constant	0.939** [0.046]	0.937** [0.047]	0.300** [0.061]	0.309** [0.067]
$\lambda$		0.903** [0.097]		0.864** [0.119]
$\rho$			0.930** [0.068]	0.920** [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 <sup>2</sup>	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.

## 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.

## **Non-exclusive licence to reproduce thesis and make thesis public**

I,

Laura Helena Kivi,

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to:
  - 1.1. reproduce, for the purpose of preservation and making available to the public, including for addition to the DSpace digital archives until expiry of the term of validity of the copyright, and
  - 1.2. make available to the public via the web environment of the University of Tartu, including via the DSpace digital archives until expiry of the term of validity of the copyright,

"Spatial dependence in European labour markets",

supervised by Tiiu Paas,

2. I am aware of the fact that the author retains these rights.
3. I certify that granting the non-exclusive licence does not infringe the intellectual property rights or rights arising from the Personal Data Protection Act.

Tartu, **24.05.2017**