SPATIAL ASPECTS OF UNEMPLOYMENT IN THE VISEGRAD-GROUP ECONOMIES

Tomáš Formánek, Roman Hušek

Abstract
Purpose of the article: Most regional macroeconomic processes may not be adequately analyzed without accounting for their spatial nature: regional distances, interactions between neighbors, spill-over effects and interdependencies. This contribution focuses on various factors ruling unemployment dynamics in the Visegrad Group countries and their major economic partners: Germany and Austria. The analysis is performed at the NUTS2 level.
Methodology/methods: Spatial econometrics is a unique tool for a broad range of quantitative analyses and evaluations. Spatial econometric models are based on geo-coded (spatially defined) data. Spatial econometrics and regional competitiveness paradigms are combined into different types of regression model specifications describing unemployment dynamics. Alternative spatial structures (i.e. neighbor definitions) are used for verification of stability in estimated model properties. Scientific aim: We aim to provide a detailed empirical evaluation of spatially determined factors of regional unemployment dynamics, along with insight into the robustness of such approach. Both conceptually and parametrically varying neighbor definitions are used to provide evidence for model evaluation. Findings: We find strong positive spatial dependence patterns in the estimated models, robust against varying neighborhood definitions. Our results strongly support the importance of regional and potentially cross-border (international) cooperation in macroeconomic policies addressing unemployment. The estimated models also underline the importance of using spatial models, by pointing out the bias in OLS-estimated models. Conclusions and limits: Spatial approach to econometric analysis provides important insight and robustness to a broad range of unemployment analyses that may be carried out using regional (spatial) data. At the same time, it should be noted that this article focuses mostly on the spatial and stability aspects of model estimation, while leaving out other interesting topics such as spill-over effects calculations as based on estimated models. Also, estimations provided in this article might benefit from spatial panel data-based methods – once data availability issues are sorted.

Keywords: Spatial econometrics, Unemployment, Regional competitiveness, Visegrad Group

JEL Classification: C23, C31, C52, E66

DOI: 10.1515/cks-2016-0007
Introduction

Spatial econometric models explicitly address the presence of spatial effects (such as economic spill-overs) when analyzing the relationships between variables using regression models and other quantitative estimation methods. Spatial models play an ever more important role in regional macroeconomic and social analyses, real estate studies, agricultural & ecological applications, epidemiology and in many other non-economic fields of research. For this type of unemployment analysis, data usually need to be geo-coded by means of the latitude/longitude geographic coordinates system\(^1\), as distances and common borders are used to estimate spatial dependencies.

It may be argued that much of the spatial effects (spatial dependencies) are attributable to omitted variable factors. However, spatial autocorrelation may be conveniently interpreted as a proxy for many real and theoretically sound, yet practically unobservable spatial effects - many spatial interactions and their dynamic features are very difficult to explicitly define and properly structure in a way that would facilitate informative and harmonized quantification. Tasks such as consistently measuring cross-border work commuting preferences, accounting for administrative/qualification employment barriers between countries, quantifying the impact of language differences, aerial distances vs. topology, etc. would inherently introduce many subjective decisions and – in practical terms – many disputable features to quantitative models. Hence, spatial models may provide a useful, interpretable and functional approach towards regional (macroeconomic) data analysis.

The remainder of this paper is structured as follows: Section one covers key topics of the spatial approach to econometrics, section two provides some additional literature review and section three focuses on regional competitiveness and unemployment dynamics. Section four discusses the empirics of spatial models being applied at the regional level NUTS2 (as defined by the Nomenclature of Territorial Units for Statistics); estimated models are interpreted and discussed in section 5. The last section and the list of references conclude our contribution.

1 Spatial autocorrelation and spatial econometric models

Before estimating spatial models, we should apply preliminary tests for spatial autocorrelation in the observed cross-sectional data. Many types of spatial autocorrelation test statistics are available (see Anselin, Rey 2010), yet Moran’s I seems to be the most widely used:

\[
I(x)_t = \left( \frac{n}{\bar{S}} \right) x_t^T W x_t (x_t^T x_t)^{-1},
\]

where \( x_t \) is the vector of \( n \) spatial observations (units) of the variable \( x \) under scrutiny at time \( t \).

\( S = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \) is the standardization factor corresponding to the sum of all elements of the spatial weights matrix \( W \). The expected value of Moran’s I under the null hypothesis of no spatial autocorrelation is: -1/(n-1). As in (Ullah, Giles 1998), \( \text{var}(I(x)_t) \) is used to calculate a z-score and test for statistical significance: whether neighbor units are more similar to one another than they would be under the \( H_0 \) of spatial randomness. The sign of Moran’s I discriminates between positive and negative spatial autocorrelation.

Spatial weights matrix \( W \) is the corner stone of spatial econometrics and, perhaps surprisingly, its construction is the most ambiguous part of the otherwise well rooted methodology of spatial model specification and estimation. \( W \) is usually calculated in a two-step approach: First, a square spatial matrix is used to define neighbors (spatially close observations) using a dummy variable technique, where each element of the symmetric spatial matrix equals 1 if the two spatial units are neighbors and 0 otherwise. Then, \( W \) is constructed by row-standardizing the spatial matrix, so that the row weights sum up to 1, while diagonal elements of \( W \) are set to zero by definition (units are not neighbors to themselves).

---

\(^1\) Geo-Data source used in this article: GISCO - Eurostat (European Commission); Administrative boundaries: © EuroGeographics.
The first step (W construction) often requires extensive geographical (polygon-based) mapping datasets and specialized software. Contiguity approach is a theoretically simple yet computationally complex rule, defining two units as neighbors if they share a common border. A generalization of this approach is based on the premise that a second order neighbor (considered as a neighbor) is the neighbor of a first order (actual) neighbor – where the maximum accepted order of neighborhood (neighborhood lag) may be set arbitrarily. Distance-based approach usually constructs the spatial matrix by defining two units as neighbors if their distance does not exceed some ad-hoc predefined threshold. This is a relatively popular approach, yet it generates “islands” (units with zero neighbors), unless the defined threshold for distance between neighbors is greater than the maximum first nearest neighbor distance. The maximum distance-based approach is less convenient for analysis of regions with uneven geographical densities – sizes of units and distances between them. Distances are measured using centroids, conveniently chosen representative positions for each unit. Depending on model focus, data availability and researcher’s individual preferences, centroids may be pure geographical center points, locations of main cities, population-base weighted positions, transportation network based, etc. Alternatively, we may apply a k-nearest neighbors (KNN) approach, where we denote a preset number of k nearest units as neighbors. This method conveniently solves for differences in areal densities (k neighbors are ensured for each unit), yet it usually leads to asymmetric spatial matrices with potentially flawed neighborhood interpretation (simple transformation algorithms for asymmetric spatial matrices are available).

The second step (W construction) usually consists of row-standardizing the binary 0/1 neighborhood indicators of the spatial matrix into matrix W so that all rows sum to unity. However, with increasing variance in units’ neighbor-count (e.g. for distance-based neighbors with uneven geographical density), this widely adopted approach suffers from allocating excessive influence to links from units with few neighbors. To overcome this drawback, sometimes the non-zero elements in W matrix are “generalized” before the row-standardization. For example, distances to neighbors are used to reflect some prior information concerning the spatial dependency processes: often we assume that spatial influence is inversely proportional to distance (linear, quadratic or other functional forms of influence decay may be used). The efficiency of any such W generalization crucially depends on the accuracy/validity of the prior information used.

The variety of available neighbor definition approaches, the choice of centroids and possible W standardization methods imply that researchers usually need to consider several different choices (spatial structure settings) in order to verify model stability and robustness. As far as spatial matrices are concerned, there usually isn’t a single right solution and researches often look for the most useful or interpretable model setup.

Once significant spatial dependence in observed data is verified, spatial regression may be used to account for such situation. Again, various model specifications and estimation methods are available. Spatial lag model is used when focusing on the analysis of spatial interactions in the dependent variable (here, the dependent variable is the one with spatial structure). A general formula for the spatial lag model and its reduced form may be written as

\[ y_t = \rho W y_t + X_t \beta + u_t, \]  
\[ (I - \rho W)y_t = X_t \beta + u_t, \]  

where \( y_t \) is the vector of all \( y_{it} \) spatial unit observations at time \( t \), \( I \) is the \( n \times n \) identity matrix, \( X_t \) is a matrix of regressors (includes the intercept element, may include lagged variables). Maximum likelihood (ML) approach is used to estimate both the spatial dependence parameter \( \rho \) and the regression coefficients \( \beta \) which are used to explain the variability in individual \( y_{it} \) observations that is not explained spatially. \( u_t \) and its elements \( u_{it} \) describe the random part of the regression model. The spatially defined portion of each dependent variable \( y_{it} \) in (2) and (3) may be expressed in terms of weighted averages of its neighbors’ values: \( W y_t \) is the spatial lag of \( y_t \) and we may write \( SpatialLag(y_{it}) = \mathbf{w}^T_i y_t \), where \( \mathbf{w}^T_i \) is a row vector: the \( i \)-th row of \( W \) matrix is used for calculation

DOI: 10.1515/cks-2016-0007

1.
(W has zeros on the diagonal and is row-standardized). Additional detailed discussion is provided in (Anselin, Rey 2010).

Even if the research is not focused on the spatial dependence and its analysis, we may still take advantage of the spatial error model to account for (correct for) the spatial nature of the observed data. In this case, we deal with spatial autocorrelation by introducing spatially correlated errors \( u_t = \lambda W u_t + \varepsilon_t \) that are justified by the presumed existence of unobservable features associated with location or with spatially defined omitted variables. By analogy to equations (2) and (3), the spatial error model and its reduced form may be expressed as

\[
y_t = X_t \beta + \lambda W u_t + \varepsilon_t, \tag{4}
\]

\[
(1 - \lambda W) y_t = (1 - \lambda W) X_t \beta + \varepsilon_t, \tag{5}
\]

where \( \lambda \) is the spatial dependence parameter and individual \( \beta \) coefficients are used to explain the variability in \( y_t \) observations that is not explained by the spatial nature of the observed data. While \( u_t \) suffers from spatial autocorrelation, \( \varepsilon_t \) is a vector of truly random elements. Estimated models (3) and (5) may be compared upon their maximized log-likelihoods. Specification test for equation (5) is available and based on the spatial common factor hypothesis that exploits the fact that model (5) may be expressed in spatial lag form (3) if spatially lagged regressors are explicitly included in \( X_t \) and if specific common factor constraints on \( \beta \)-coefficients hold. Spatial lag of the \( k \)-th regressor is defined as:

\[
SpatialLag(x_{kt}) = w^T \text{ } x_{kt}.
\]

Additional information on spatial models such as (2) to (5) and further model selection and validation topics are discussed in (Anselin, Rey 2010) who also provide complementary literature references.

As the \( n \times n \) dimension of the W matrix is determined by the number of spatial units, computational limits for the methods described here exist. However, econometric models comprised of, say, 500 and more units may be easily handled using ordinary PC configurations and the freely available R software (http://www.r-project.org), especially if spatial matrices are sparse (each unit has a relatively limited amount of neighbors).

2 Spatial econometrics - brief literature review

Moran (1950) and Geary (1954) are often cited as the founding fathers of spatial econometrics, yet the actual framework for contemporary applied spatial analysis was provided through introducing a flexible spatial weights specification, as in Cliff and Ord (1969, 1981 and subsequent publications). Fujita et al. (1999) provide a comprehensive theoretical background of the so called Economic Geography, which emphasizes the importance of spatial spillovers between economies and focuses on economic convergence models. Practical aspects of W matrix construction, along with model specification and estimation topics within the R software environment are covered by Bivand (2015).

Also, spatial regression models have been extended to encompass the time dynamics aspects of the geo-coded variables and their interdependencies. Elhorst (2014) provides an advanced yet concise textbook on modern spatial panel data methods, while Baltagi et al. (2012) and Fingleton (2014) published papers with applied analyses and forecasts based on spatial panel data models; they also provide many additional references to relevant literature.

3 Regional competitiveness and unemployment

At the microeconomic (firm) level, the notion of competitiveness is fairly clear and based on the individual firm’s capacity to compete, grow and make profit through products or services that meet market requirements in terms of price and quality (other relevant market factors may play a role). However, at the macro level, the concept of country-based or regional competitiveness and its definitions is less clear, even though competitiveness is frequently presented as the main goal of economic policy actions. Using the political paradigm, competitiveness is usually perceived as some combination of productivity, employment (unemployment) rate, living standards, foreign direct investment (FDI) attractiveness, etc. However, individual and collective political preferences, country-specific short term fluctuations in key macroeconomic indicators, general public attitudes and other rather
amorphous influences may play a potentially significant role in definition of competitiveness, leading to ambiguous interpretations and possible misunderstandings.

Some authors (e.g., Krugman, 1994) consider the concept of national and regional competitiveness to be mostly meaningless and potentially dangerous, arguing that the analogy between nations (regions) and firms is inherently flawed: unsuccessful firms are ultimately supposed to go out of business whereas no equivalent situation for a nation or region should occur as far as developed countries are concerned. Martin (2004) provides detailed and structured discussion of macroeconomic competitiveness aspects.

This contribution focuses on spatially defined macroeconomic variables that are consistent with the mainstream “consensus view” (see Martin, 2004) of regional competitiveness theory, where successful regional macroeconomic performance may be assessed in terms of living standard dynamics (i.e. GDP per capita in absolute or relative terms) and in terms of unemployment and its dynamics. Both GDP per capita and unemployment are often cited as constituent parts of quantitative models of regional competitiveness, perhaps along with some convenient technological advantage indicator (e.g. high-tech sector labor force proportion). For in-depth discussion, please refer to Gauselmann, Marek, Angenendt (2011).

Graph 1 NUTS2-level unemployment rates, comparison of years 2005 & 2014

Graph 1 shows two choropleths (i.e. info-map graphs) providing spatially defined and time-specific comparisons of unemployment in Visegrad Group countries, Austria and Germany. From Graph 1, we may observe prominent spatial autocorrelation patterns – observations are spatially clustered for both years displayed. Individual choropleths may be briefly discussed as follows: In 2005, DOI: 10.1515/cks-2016-0007

1.
the unemployment data show a clear difference in labor-market situation between former ‘Western EU’ (Austria and ‘West Germany’ i.e. Federal Republic of Germany) and ‘Eastern EU’ (all remaining) regions. From the 2014 unemployment choropleth, we may see that much of the clear distinction between unemployment rates in the “old” and “new” EU regions is dissipated between 2005 and 2014 (partly due to an overall decline in unemployment), yet the differences remain somewhat visible. The 2005 unemployment data is provided for readers’ convenience and reference only, the empirical part of this article is based on recently observed data.

Graph 2 consists of two choropleths for two additional competitiveness-related variables: the percentage rate of employment in high tech sector and GDP per capita. The percentage rate of employment in high tech sector shows a strong positive correlation with the GPD indicator, as well as a strong negative correlation with unemployment (see Graph 1). The GDP per capita variable suggests a lack of wealth-convergence between Germany and Austria on one side and former ‘Eastern’ countries on the other side. The only distinctive exception is the SK01 NUTS2 region (Bratislava).

Country-wise situation considered, Austria exhibits a strong stability in observations, as its color-coded data hardly change in any of the choropleths. The Slovak Republic shows a very prominent core-periphery pattern, as all favorable values (low unemployment, high GPD per capita indicator, strong high-tech sector) deteriorate with distance from SK01 (Bratislava) eastwards, i.e. towards the border with Ukraine (EU border). The Czech Republic manifests analogous core-periphery behavior with unemployment indicators worsening with distance from CZ01 (Prague). However, there is one important exemption from this pattern: the well-performing CZ03 (Southwest region) apparently
benefits from bordering with high-performing Bavarian and Austrian regions. Interestingly, no such positive effect may be observed for CZ04 (Northwest region), which shares most of its borders with former East German regions. Similar, yet less prominent core-periphery behavior may also be observed in data from Poland and Hungary. Germany seems to have at least two highly competitive centers (Bavaria and Hesse NUTS1 regions) that are both located in the former ‘West Germany’, while East-West differences are becoming less distinctive over time.

4 Empirical analysis and interpretation of results

Using the quantitative and economic paradigms outlined in previous sections, selected regional competitiveness indicators may be assembled into a relatively simple, yet useful, informative and theoretically well-defined spatial model explaining unemployment dynamics in terms of per capita productivity and using a convenient technological advantage indicator (high-tech sector labor force proportion).

4.1 Data

Even though Eurostat (http://ec.europa.eu/eurostat) has made a great progress in harmonization and availability of spatial macroeconomic and socio-economic data at ‘lower’ levels of NUTS aggregation (NUTS2, NUTS3) over the last years, data availability issues are present in most data series relevant for this study. Unfortunately, this impedes the use of panel data methods and somewhat limits the selection of other suitable estimation approaches. At the NUTS2 level, there is a great likelihood of missing observations across regions during each year included in the dataset (annual data are used, as relevant quarterly observations are completely unavailable for most variables). The set of 82 geo-coded cross sectional observations at the NUTS2 level (2010 NUTS specification) is comprised of 8 regions in the Czech Republic, 4 in the Slovak Republic, 9 in Austria, 38 in Germany (of those, 8 are former East-German regions; 9 including the unified Berlin), 16 in Poland and 7 in Hungary. R-software was used to extract GDP per capita, unemployment and high-tech employment data from Eurostat. Specifically, the following databases are used (identification by table code): “nama_10r_2gdp”, “lfst_r_lfu3rt” and “htec_emp_reg2”. Although observations in some spatial data series may cover the period from 1999 to 2016, actual data availability issues and temporal relevance considerations lead to focusing the research on the year 2014 as suitable and complete datasets for 2015 are not yet completely available at the NUTS2 level. Finally, it should be noted that we had to impute missing 2014 HighTech employment rate data for the PL62 region (Warminski-Mazurskie): we use a value of 1.2% - the same value was observed in this region in both 2012 and 2013.

4.2 Model specification

The application part is based on regressing regional unemployment on GDP (logs of GDP per capita) and a convenient technological advantage measure (relative employment in high-tech industries). Using the general specification of spatial lag and spatial error models (2) and (4), equations (6) and (7) represent the chosen models used for estimation and subsequent analysis of regional unemployment dynamics:

\[ Unem_{it} = \rho w_i^T Unem_t + \beta_0 + \beta_1 \log(GDP_{it}) + \beta_2 TechEmp_{it} + \beta_3 HiUnCl_{it} + u_{it}, \]

\[ Unemp_{it} = \beta_0 + \beta_1 \log(GDP_{it}) + \beta_2 TechEmp_{it} + \beta_3 HiUnCl_{it} + \lambda w_i^T u_t + \varepsilon_{it}, \]

where \( Unem_i \) is the general rate of unemployment for a given NUTS2 region \( i \) at time \( t \) (2014), \( Unem_t \) is a vector of all \( Unem \) observations for the year 2014. \( GDP_{it} \) is region’s GDP per capita (current prices). The expression \( \log(GDP_{it}) \) may be interpreted in terms of relative (percentage) changes. The dynamic interpretation is consistent with neo-classical macroeconomics, as described by Williams (2005) and others. The competitiveness-related variable \( TechEmp_i \) describes the percentage of employees working in the “high-tech industry” (NACE r.2 code HTC) in each region. \( HiUnCl_i \) is a dummy variable based on Ord’s \( G^* \) (local \( G \) as in Ord, Getis 1995) that discerns local clusters (hotspots) of high values of the variable being analyzed – as we searched for high unemployment clusters,

DOI: 10.1515/cks-2016-0007
we found a single cluster, containing the following NUTS2 regions: HU10, HU31, PL21, PL22, PL32, PL33, SK03 and SK04 (generally speaking this cluster consist of units close to or bordering with Ukraine).

4.3 Model estimation

At the 5% significance level and across varying neighbor structure definitions, Moran’s I as per equation (1) test results provide a strong evidence against the null hypothesis of spatial randomness in all observed variables as described in the previous paragraph. Due to space limitations, individual Moran’s I results are omitted here. All relevant statistics, figures, estimated models and tests mentioned here are available from the authors upon request. As Moran’s I results suggest strong and positive spatial autocorrelation for all observed data series. OLS method is not an appropriate estimator. However, it may provide a base-reference for comparison against more sophisticated models.

Table 1 shows the core estimation output, where column (a) is obtained by OLS with no spatially defined information included (model without unemployment clustering) and column (b) is obtained using OLS (unemployment hotspots included). Columns (c) to (g) provide a representative set of estimation results for various spatial dependence specifications. Overall, Table 1 demonstrates both the superiority of spatial models against OLS as well as a reasonable robustness of spatial models against changes in neighborhood definition and spatial lag / spatial error model specification.

Results in columns (a) and (b) were verified against heteroscedasticity and multicollinearity. Specialized Lagrange multiplier test statistics for spatial autocorrelation were used to decide whether spatial lag (6) or spatial error model (7) should be used for estimation (see Anselin, Rey 2010). At the 5% significance level, spatial lag model is supported by the observed data across all selected spatial lag (6) or spatial error model (7) should be used for estimation (see Anselin, Rey 2010). At the 5% significance level and across varying neighbor structure definitions, Moran’s I as per equation (1) test results provide a strong evidence against the null hypothesis of spatial randomness in all observed variables as described in the previous paragraph. Due to space limitations, individual Moran’s I results are omitted here. All relevant statistics, figures, estimated models and tests mentioned here are available from the authors upon request. As Moran’s I results suggest strong and positive spatial autocorrelation for all observed data series. OLS method is not an appropriate estimator. However, it may provide a base-reference for comparison against more sophisticated models.

Table 1 Outputs from alternative model estimation methods

| Coefficient est. (standard error) [p-val., z-val.] | (a) OLS (b) OLS (c) Spatial lag model (d) Spatial lag model (e) Spatial error model (f) Spatial error model (g) Spatial error model | | | |
| (var. name) | (unemp-) spatial inf. clustering used) | (binary, 240km dist.) | (binary, 240km dist.) | (KNN method, k=4) | (KNN method, k=4) | (KNN method, k=13) |
| log(GDP) | -3.231* (0.408) | -2.580* (0.409) | -0.665 (0.415) | -2.090* (0.402) | -0.990* (0.398) | -1.164* (0.652) | -0.501 (0.411) |
| TechEmp | -0.012 (0.152) | -0.071 (0.140) | -0.194* (0.117) | -0.056 (0.131) | -0.165 (0.113) | -0.144 (0.133) | -0.186 (0.116) |
| HiUnCl | --- (0.817) | --- (0.702) | 1.922* (0.839) | 2.159* (0.692) | 1.416* (0.949) | 1.267* (0.680) | 2.144* (0.680) |
| P | --- --- 0.705* (0.107) 0.309* (0.108) 0.606* (0.096) --- --- 0.759* (0.100) |
| A | --- --- --- --- --- 0.707* (0.086) --- --- |
| Log likelihood | -178.69 -171.24 -160.63 -168.07 -159.03 -164.24 -159.16 |
| AIC | 365.38 352.47 333.26 348.13 330.05 340.47 330.33 |
BIC  375.01  364.51  347.70  362.57  344.49  354.91  344.77
Wald test  [p-value]  37.91* (0.000)  35.12* (0.000)  18.18* (0.000)  47.03* (0.000)  18.62* (0.000)  11.32* (0.010)  19.16
$R^2 = 1 - \frac{SSR}{SST}$  0.490  0.575  0.692  0.611  0.714  0.690  0.702

Source: self prepared using R software

Basically, any spatially-augmented estimation of the model is superior to the OLS-based model estimates as measured by the maximized log-likelihood (LL) statistics, Akaike and Bayes information criteria (AIC, BIC), etc. The $KNN$-based spatial lag model with $k = 4$ as shown in column (e) leads to the best results as far as all model-evaluation criteria are concerned. However, the performance of the model as depicted in column (c) is almost equally good. Given the results of the Anselin-Rey test statistics (Anselin, Rey 2010), the $KNN$ method for $k = 4$ also provides an interesting opportunity to compare coefficients from spatial lag and spatial error models (e) and (f): both specifications confirm strong spatial dependence and lead to reasonably similar $\beta$ coefficients. Maximum neighbor distance threshold methods (c) and (d) provide other effective ways to define $W$ and estimate a spatial lag model. For column (d), we assume that spatial dependence decays quadratically (along the square of distance).

4.4 Estimation robustness - stability of results

Given the multiple possible approaches to spatial matrix and $W$ matrix construction and the resulting ambiguity potential, specification robustness was tested against changes in neighborhood definition for the spatial lag model (6): Multiple estimations of model (6) were performed with the dataset available, while varying the spatial weights matrix $W$. A representative robustness analysis output is shown in Graph 3 and it may be briefly summarized as follows: Starting from a very sparse $W$ matrix constructed using maximum neighbor distance threshold set to 160 km (lower distances generate “islands” that are incompatible with ML estimation of spatial models), neighbor thresholds are increased and new $W$ matrices generated by iterations of 10 km, up to a maximum neighbor distance of 1.000 km. Hence, a total of 85 alternative spatial structure-specifications of model (6) are produced, estimated and summarized in figure 2. Results obtained using distance thresholds beyond 1.000 km provide no significant improvement over OLS: $\text{var} \left[ \hat{\beta}_\text{Unem} \right]$ falls quickly with the number of neighbors approaching its maximum value of 81. Graph 3 shows the AICs, $\beta$ and $\rho$ coefficients with their asymptotic (+/-) one standard error bands.
Relative model (estimate) instability at the lower end of the distance threshold interval should not be viewed as puzzling: 160 km is the minimum distance that avoids islands, but such spatial structure is not realistic (providing very few neighbors for the spill-over effects to take place on any observed region). High distance thresholds are not economically realistic either, as they are not consistent with prevalent regional interactions for unemployment. For readers’ convenience, Table 1, column (c) results are highlighted by a dotted vertical line in Graph 3: it corresponds to the maximum distance threshold of 240 km – where the best (lowest) AIC values are obtained.

As we evaluated the fitted values of spatial models shown in columns (c) to (g) of Table 1, we may conclude that the specifications used provide a reasonable basis for the usage of the spatial model for prediction purposes for all regions considered. Specifically, the results do not show any significant bias towards, say, accurately predicting unemployment in German regions (38 of the total 82 spatial units) at the expense of other countries (i.e. NUTS0 regions). A significant proportion of this behavior may be attributed to the heterogeneity of German regions, where 8 NUTS2 units belong to the former “East Germany” (9 regions including the unified Berlin). Overall, the dataset contains a favorable heterogeneous mix of NUTS2 regions observed at diverse “performance” levels, leading to satisfactory predictive properties of the model.

Discussion and interpretation of the results

Once spatial dependence in the data is properly addressed, the regression model may serve as a tool for discerning the influence of geographically defined conditions from factors that may be - at least potentially - influenced by macroeconomic policies such as regionally focused fiscal stimuli, establishing or subsidizing technological parks, applied research funding, tax incentives, etc.
Across all alternative model specifications (a) to (g) from Table 1, we may see that a marginal improvement in any of the competitiveness indicators (model regressors) leads to an expected decrease in unit’s unemployment rate. However, many $\beta$ coefficients in spatially augmented models are substantially attenuated (“scaled down” towards zero) as compared to the OLS model coefficients in columns (a) and (b). Along with strong positive spatial dependence (high and significant coefficients $\rho$ and $\lambda$), such results have two main interpretations: they strongly support the importance of regional and potentially cross-border (international) cooperation in macroeconomic policies addressing unemployment, plus they underline the importance of spatial models, by pointing out the bias in OLS-estimated expected effects of changes in economic policies and the expected results of other ad-hoc events bearing the potential of influencing unemployment.

Conclusions
Spatial econometric models provide a useful estimation framework that allows for improved analyses of regional macroeconomic data. Spatial models have a unique ability to discern between geographical determination (spatial autocorrelation, spill-over effects) and the influence of relevant macroeconomic variables, many of which may be subject to or directly controlled by economic policy actions as undertaken by the central authorities at different levels.

Our results strongly support the importance of regional and potentially cross-border cooperation in macroeconomic policies addressing unemployment. Robust spatial dependencies were identified in observed data - spatial autocorrelation tests are consistently significant for a very wide range of neighborhood definitions. This study also provides relevant implications towards the analyses of dynamics in other macro-indicators of competitiveness. As regional observed data display strong spatial autocorrelation, either spatial lag or spatial error models should be considered while assessing macroeconomic dynamics, its properties and relevant implications towards economic policy actions.

References

DOI: 10.1515/cks-2016-0007


**Author(s) contact(s)**

**Ing. Tomáš Formánek, Ph.D.,** University of Economics, Prague, dept.of Econometrics, W. Churchill sq. 4, Prague, 130 67; phone: +420 224095423; email: formanek@vse.cz

**Prof. Ing. Roman Hušek, CSc.,** University of Economics, Prague, dept.of Econometrics, W. Churchill sq. 4, Prague, 130 67; phone: +420 224095423; email: husek@vse.cz

**Acknowledgement:** Supported by the grant No. IGA F4/73/2016, Faculty of Informatics and Statistics, University of Economics, Prague.