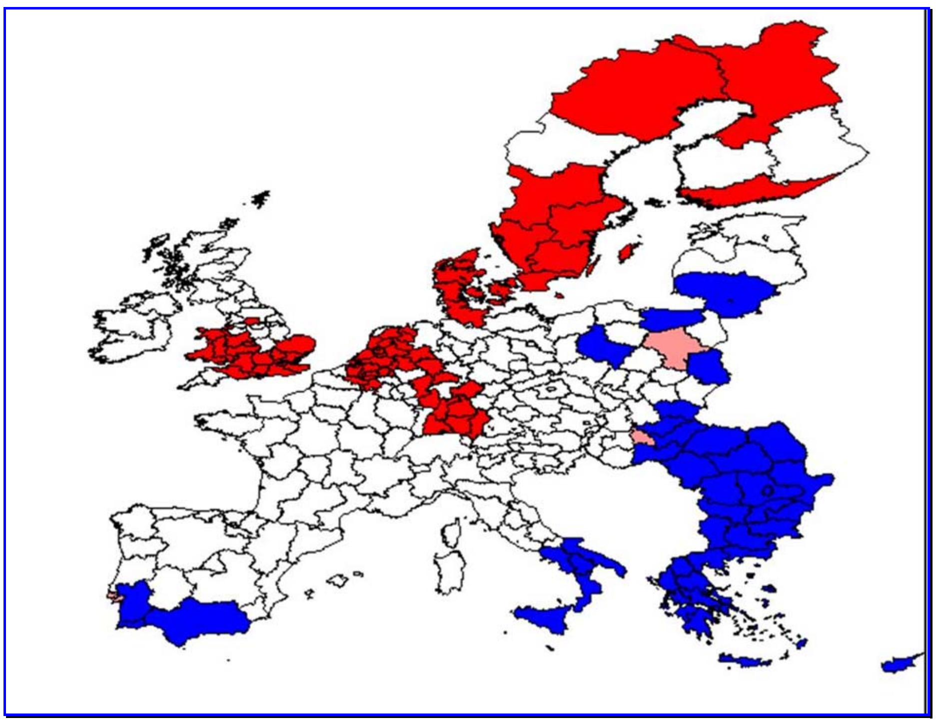
JRC Scientific and Technical Reports

RCI 2010:
Some in-depth analysis

Paola Annoni and Kornelia Kozovska



EUR 24703 EN - 2011

JRC EUROPEAN COMMISSION
Institute for the Protection and Security of the Citizen

The mission of the JRC-IPSC is to provide research results and to support EU policy-makers in their effort towards global security and towards protection of European citizens from accidents, deliberate attacks, fraud and illegal actions against EU policies.

European Commission
Joint Research Centre
Institute for the Protection and Security of the Citizen

Contact information
Address: Econometrics and Applied Statistics Unit, Via E. Fermi 2749, Ispra (VA), Italy
E-mail: paola.annoni@jrc.ec.europa.eu, kornelia.kozovska@jrc.ec.europa.eu
Tel.: +39 0332 78 6448
Fax: +39 0332 78 5733

http://ipsc.jrc.ec.europa.eu/
http://www.jrc.ec.europa.eu/

JRC 62665

EUR 24703 EN
ISBN 978-92-79-19078-0
ISSN 1018-5593
doi:10.2788/28476

Luxembourg: Publications Office of the European Union

Printed in Italy

Table of contents

Executive summary

This document is the final delivery of the two-year joint project DG Joint Research Centre and DG Regional Policy on the measurement of the level of regional competitiveness, launched in November 2008. Within this project, the European Commission has recently published the first edition of the Regional Competitiveness Index (RCI) (Annoni and Kozovska, 2010). The index provides a tool to improve the understanding of competitiveness at the regional level by showing the strengths and weaknesses of each of the European regions at the NUTS2 level in a number of dimensions related to competitiveness. The analysis offered by the first edition of the RCI is a snapshot of regional competitiveness as it is in 2010 and is based upon data mostly spanning between 2007 and 2009. The present document takes a step further and offers a two-fold analysis based on the RCI indices: an exploratory spatial data analysis and an analysis of possible relationships between exogenous indicators and the RCI index and sub-indices.

The exploratory spatial data analysis shows the existence of spatial dependence among EU regions, with different patterns for different areas within the EU. This can be taken as an indication for the existence of spatial externalities among regions and, when observed for high performing regions, as evidence, or better, as necessary condition for spillover effects. The Moran's global index of spatial autocorrelation shows that there exists spatial autocorrelation for the RCI index as well as the three sub-indices – basic, efficiency and innovation. Local clusters of low RCI values, as evidenced by the analysis of Local Indicators of Spatial Association (LISA), are located in Bulgaria, Romania, Greece and Cyprus, South-Eastern regions of Hungary and Slovakia, Southern part of Italy, Portugal and Spain. These areas show significant results for low-low clusters, meaning that regions with low RCI scores are surrounded by low scoring regions. On the other hand, local clusters of high RCI values are found in regions in the Netherlands, parts of Germany, Belgium, Denmark, the southern part of United Kingdom, Finland and Sweden.

Overall, LISA analysis allowed us to distinguish between two sub-areas in the EU: group A which comprises regions with high RCI performance surrounded by regions with similar strong competitive performance (these regions are located in the following countries: AT-BE-DE-DK-FI-IE-LU-NL-SE-SI-UK) and group B, comprising low-performing regions surrounded by low RCI performers (these regions are located in the following countries: BG-CZ-EE-GR-HU-LT-LV-PL-RO-SK).

The analysis has been extended to better explore the structure of spatial autocorrelation within the two main sub-areas – A and B - of low-low and high-high clusters as detected by LISA. The analysis of sub-area B is meant to further investigate the possible presence of 'negative' spillover effects where low performing regions negatively affect their neighbors.

Spatial autocorrelation structure is investigated by using variogram analysis, a tool typical of Kriging for describing spatial dependences (Cressie, 1984). Variogram analysis provides as additional information the 'range of action' of spatial dependence, which is the maximum distance beyond which the correlation can be considered null. Variogram analysis is carried out using three different distances between region centroids: Euclidean distance, distance along the road (ferry) network and the travel time distance. Results indicate the existence of a clear structure of correlation for the sub-area A of high-high clusters. In this area the range of auto-correlation is between 300-500 km for Euclidean and road distance, while in terms of travel-time distance the estimated range is about 150-200 minutes. Variogram analysis cannot estimate a range for the sub-area of low-low clusters, sub-area B. This area seems to be characterized mostly by low performing regions with some rare and sparse picks of relatively higher performers (some capital regions).

With regards to the analysis of possible relationships between exogenous indicators and RCI index and sub-indices, we have looked at bivariate correlations with five exogenous indicators (population change in the period 2001-2007; natural population change in the period 2001-2007; net migration in the period 2001-2007; share of population which live in Large Urban Zones, LUZ; GDP growth average 2000-2007) for all EU NUTS 2 regions as well as for two sub-areas as identified by the ESDA analysis. We find that the number of significant results from the correlation analysis increases when we distinguish between the sub-areas.

The share of population living in LUZ is always positively associated to the three RCI indices (total, efficiency and innovation), with particularly high values for countries in sub-area B. In these regions living in high density areas (large cities) means having higher levels of competitiveness. Focusing on countries in sub-area A, where spatial auto-correlation analysis highlighted clusters of highly performing regions, all the indicators but net migration and GDP growth average are positively correlated with RCI indices The analysis for regions in sub-area B shows a positive correlation for all the exogenous indicators and almost all the RCI indices (with the only exception of the efficiency sub-index and natural population change).

In general, our results show that population dynamics and demographic trends are highly relevant for territorial competitiveness while the relationship with GDP growth remains ambiguous. Two critical issues arise here: first, RCI 2010 covers a lag of time which comprises the 2008 economic and financial crisis; second, the relationship between competitiveness and growth is in general difficult to understand. A recent example is the Trade Performance Index, jointly developed by UNCTAD (United Nations Conference on Trade and Development) and WTO (World Trade Organization), where high performing countries are those where GDP growth was the lowest in the last ten years. For these reasons, an in-depth analysis of the relationship between territorial competitiveness and economic growth indicators would require a separate and extensive research which goes beyond the scope of this study.

Introduction

The regional competitiveness index recently developed by the European Commission (Annoni and Kozovska, 2010) provides a tool to improve the understanding of competitiveness at the regional level. The index shows the strengths and weaknesses of each of the European regions at the NUTS2 level and covers a wide range of issues related to competitiveness. The analysis is a snapshot of regional competitiveness as it is in 2010 and is based upon data mostly spanning between 2007 and 2009.

The present document provides a two-fold analysis based on the RCI indices: an explorative spatial data analysis and an analysis of possible relationships between exogenous indicators and the RCI score and sub-scores.

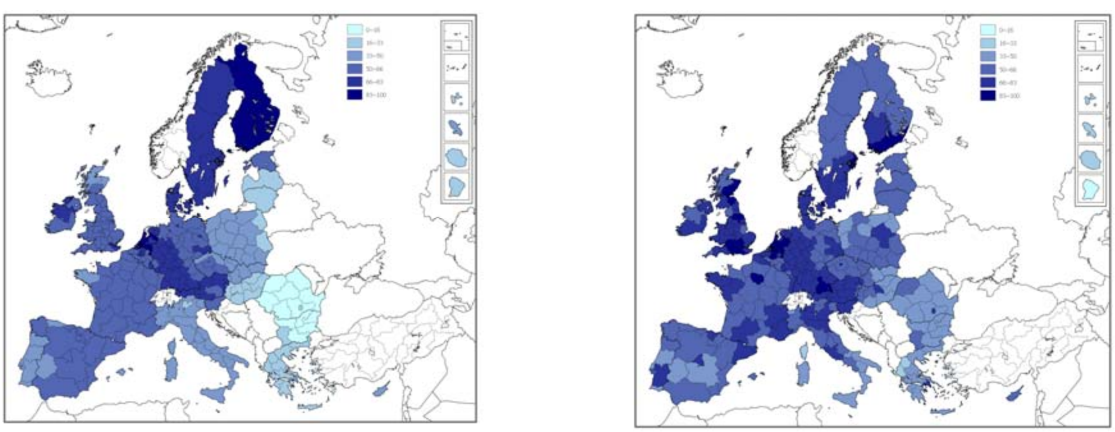
The spatial data analysis explores the structure, if any, of spatial autocorrelation of the RCI scores with the final aim of detecting clusters of high or low performers among the European regions. This is the first step in the assessment of possible spill-over effects of competitiveness. The analysis of relationships with exogenous indicators, where 'exogenous' is understood as indicators which have not been directly included in the RCI computation, may help in finding out possible drivers of competitiveness.

1 The analysis of spatial auto-correlation

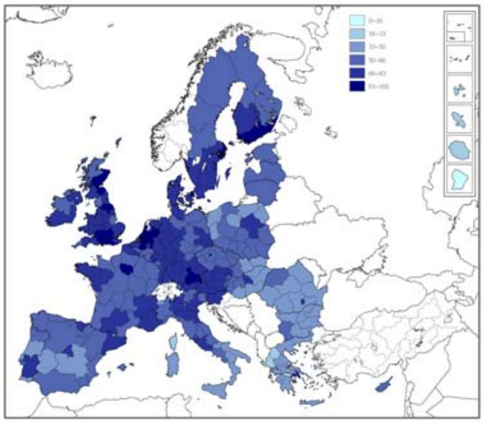
By its nature competitiveness is not bound by administrative borders, being a result of a number of factors which interact among each other as well as with the surrounding environment. Trade between regions, labor mobility, technology and knowledge diffusion and regional externalities in general are a source of geographical dependence among regions. Thus, it is quite natural to assume that competitiveness levels of different EU NUTS 2 regions influence and are influenced by the performance of their surrounding regions, giving rise to spill-over effects. Given the institutional and economic set up of the European Union, such interactions are not necessarily limited to regions within the same country but could very well exist among bordering regions from different countries.

The concept of spillovers has been widely used in economic literature to describe externalities generated by a number of processes, including geographical proximity, and concerning productivity, knowledge, innovation among others. The measurement of spillovers is dependent upon the type of process examined. At the regional level, spillovers have been studied largely in the framework of regional convergence (see Islam, 2003; Magrini, 2004 for surveys on the topic) of GDP levels. Evidence for the existence of regional spillovers has been analyzed through several types of convergence processes - catching-up in per capita income levels, usually estimated by regressing growth rates of GDP on initial levels (δ-convergence), decline in the cross-sectional dispersion, measured as the sample variance, of per capita incomes (σ-convergence) or changes in the rankings of relative per capita income (g-convergence) (see e.g. Baumont et al, 2003; Dall'Erba and Le Gallo, 2008). Reference is made to economies being similar in structural characteristics and converging within groups depending upon initial conditions or other spatial or a-spatial attributes (Ramajo et al., 2008). Many empirical studies do not include the spatial aspect of regional data in their estimation and numerous critiques have been made on the methodological aspects of studying spillovers through such convergence methodologies, especially at the regional level (e.g. Ramajo et al., 2008). Furthermore, when looking at the different types of possible spillovers (e.g. knowledge, technology, innovation, productivity) and the level (e.g. sectoral and/or regional), the possible techniques for estimating the spillover effect change according to the set-up of interest.

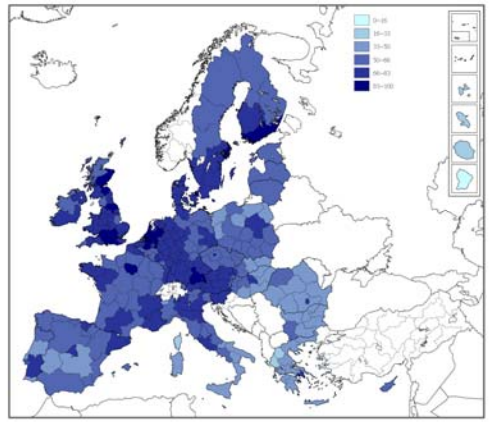
In the RCI case, the simple qualitative analysis of the maps of the total score and the three sub-indexes (Annoni and Kosovska, 2010), clearly shows that the spatial distribution of the competitiveness score is not homogeneous across EU regions (Figure 1). Apart from the map of the basic sub-index which is by construction composed by a majority of pillars at the country level, a concentration of highly performing regions (dark blue color) can be seen in the area including the following countries (clockwise from south-est): Slovenia (SI), Austria (AT), Germany (DE), Luxemburg (LU), Belgium (BE), United Kingdom (UK), Ireland (IE), The Netherlands (NL), Denmark (DK), Sweden (SE) and Finland (FI). This area may exhibit spillover effects which are here understood as strong regional economies positively influencing neighboring economies. On the other side, a spatial cluster of low values (light blue color) is detected in the area including the countries: Greece (GR), Bulgaria (BG), Romania (RO), Hungary (HU), Slovak Republic (SK), Check Republic (CZ), Poland (PL), Lithuania (LT), Latvia (LV) and Estonia (EE). This area shows indication of spatial autocorrelation (low values are 'close' to low values), as demonstrated by the ESDA analysis discussed in the following section. Such pattern does not identify potential presence of spillover effects but rather lack of such or a negative one (where low performer regions are surrounded by low performers). Other smaller areas seem to show a more heterogeneous situation with a mix of high, intermediate and low performing regions. This is the case of Spain (ES) and Portugal (PT), France (FR) and Italy (IT).



a) Map of the RCI-basic sub-index



b) Map of the RCI-efficiency sub-index



c) Map of the RCI-innovation sub-index

d) Map of the RCI final index

Figure 1: Maps of RCI index and sub-indexes

This qualitative analysis supports further investigation of the structure of spatial auto-correlation of the RCI index and the sub-indexes in search for a more quantitative assessment of spillover effects.

To analyze the spatial correlation of RCI we have opted for applying exploratory spatial data analysis (ESDA) techniques and variogram analysis in order to answer the following questions deeply interrelated with each other:

- Does spatial dependence exist for RCI?
- How much does proximity matter in the distribution of the RCI scores?
- Is there a tendency for regions with similar scores to be found close together and dissimilar ones apart?
- If spatial dependence exists, how far is it spread? What is its “range of action”?

In answering these questions our basic assumption is that the existence of spatial dependence can be taken as an indication for the existence of spatial externalities among regions and, when observed for high performing regions, as evidence, or better, as necessary condition for spillover effects. Hence, one can get an insight into spillover effects by characterizing the spatial autocorrelation structure of RCI.

1.1 Exploratory spatial data analysis - ESDA

Spatial interaction among regions and the potential presence of regional spillovers in competitiveness can be evaluated using ESDA. ESDA is a subset of exploratory data analysis (EDA) focusing on characteristics of spatial data, specifically related to spatial autocorrelation and spatial heterogeneity (Anselin et al., 2007, Anselin, 1999, Cressie, 1984, Haining, 2003). It comprises techniques for exploring spatial data such as visualizing spatial distributions, summarizing spatial properties of the data, detecting spatial patterns in data, identifying atypical locations or spatial outliers, patterns of spatial association, clusters or hot spots. ESDA compares the observed pattern in the data to one in which space is irrelevant and the spatial pattern, spatial structure, or form for the spatial dependence are derived from the data only.[1]

1.1.1 Data Visualization

The simplest and most institutive exploratory analysis is the visualization of the variable under examination on a so called ‘chloromap’. Standard chloromaps are the percentile and standard deviation maps. In a *Percentile Map*, the data are sorted and grouped in categories in order to accentuate the extreme values. A *Standard Deviation Map* groups observations according to where their values fall on a standardized range, expressed as standard deviation units away from the mean.

In the RCI case we analyse the final index RCI and the three sub-indices related to the three groups of pillars: RCI_basic, RCI_efficiency and RCI_innovation. For each of them percentile and standard deviation maps are provided.

In a standard deviation map, the variable under analysis is transformed into standardized scores (*Z* scores). This transformation puts all the scores in each distribution into the same scale where the unit of measurement is the standard deviation. In our analysis regions are classified into six classes: (μ-3σ), (μ-2σ), (μ-σ), (μ +σ), (μ+2σ), (μ+3σ), where μ is the overall arithmetic mean across all the regions and σ the sample standard deviation. Assuming a standard normal distribution for the RCI scores, which is almost

[1] All spatial analysis has been carried out using the GeoDa software package (Anselin et al. 2006) and the Matlab function ‘variogram’. French overseas territories (FR91, FR92, FR93 and FR94) have been excluded from the analysis.

the case as they have been computed as weighted averages of transformed and standardized indicators (Annoni and Kozovska, 2010), 68% of the values are expected to fall in the interval [(μ-σ); (μ+σ)] and about 95% percent in the interval [(μ-2σ); (μ+2σ)]. This means that countries with scores outside the interval [(μ-2σ); (μ+2σ)] are very low/high performers as they count for less than 5% of the score distribution. Standard deviation maps can give a quick glance of extreme cases as well as 'average' cases. It is also possible to pinpoint whether extreme cases are clustered.

As we can see from the RCI percentile map in Figure 2 regions in Bulgaria, Romania and Greece show high concentration of low values in the lowest percentiles while regions in the Netherlands and parts of Germany have concentration of high values in the highest percentiles. The situation in Spain, Italy and France is very heterogenous. United Kingdom regions also show different performances but mostly concentrated on the high percentiles while Central Eastern European regions are concentrated on the lower ones. Figure 3 shows that low and very low performing regions belong to a sector which goes from Portugal and Spain to Romania and Bulgaria, via southern Italy and Greece. Excellence areas are located in Germany and Benelux, United Kingdom and the southern part of Scandinavia.

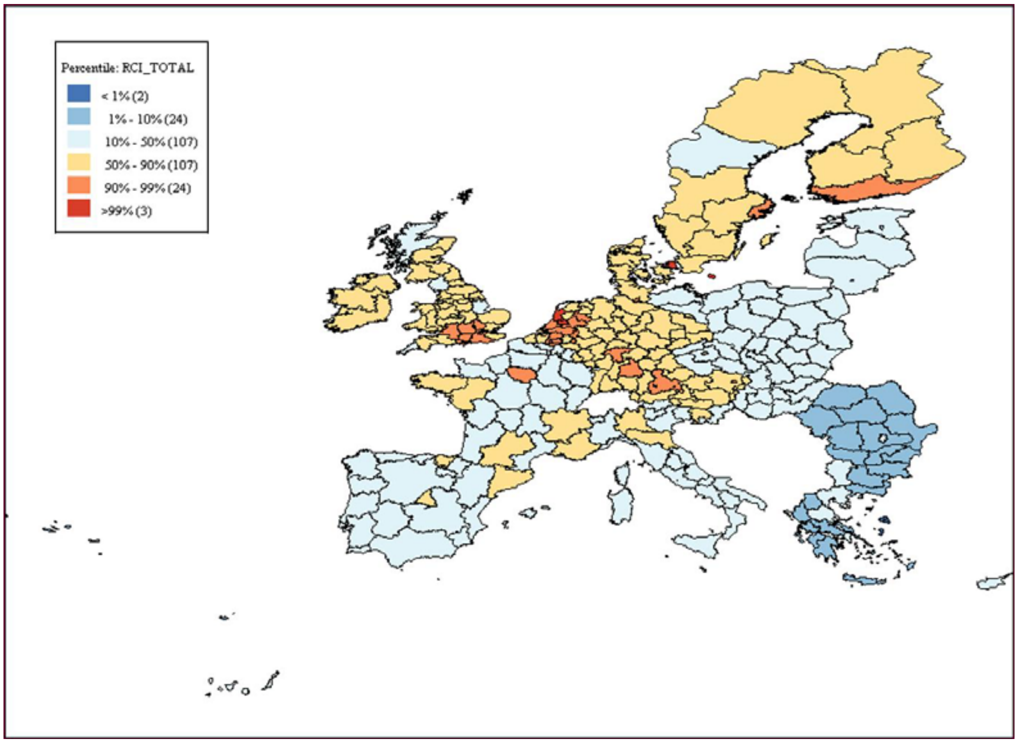


Figure 2: RCI percentile map

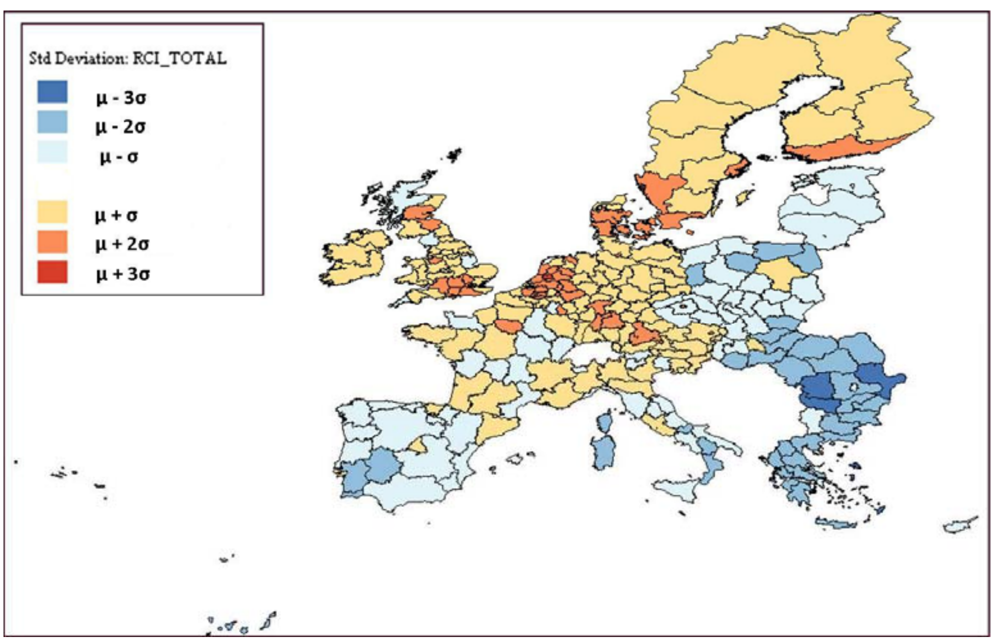


Figure 3: RCI standard deviation map

Percentile and standard deviation maps are also plotted for the three RCI sub-indices separately (Figure 4 and Figure 5 refer to the basic sub-index; Figure 6 and Figure 7 refer to the efficiency sub-index; Figure 8 and Figure 9 refer to the innovation sub-index). The larger areas of homogeneous color of the maps for the basic sub-index are due to the fact that three out of five dimensions included in the sub-index are measured at the country level.

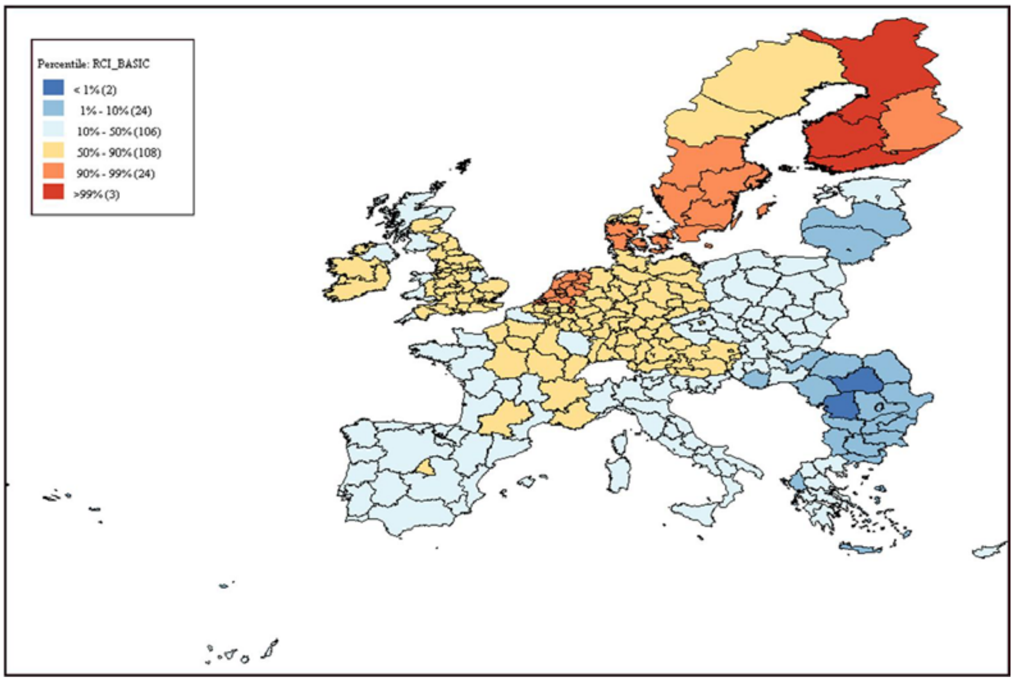


Figure 4: RCI-basic percentile map

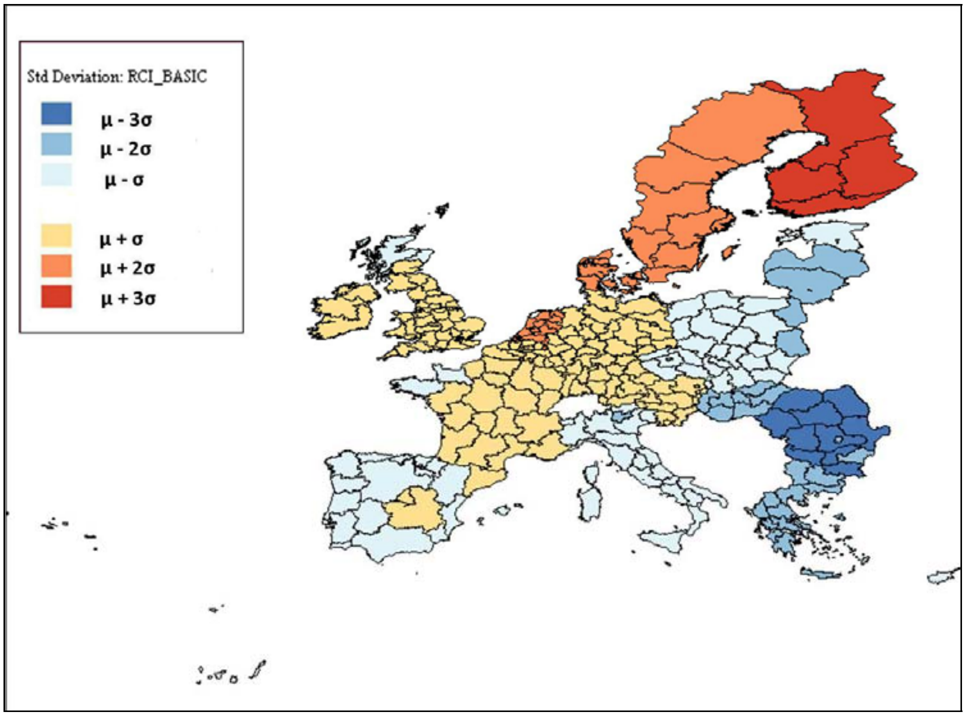


Figure 5: RCI-basic standard deviation map

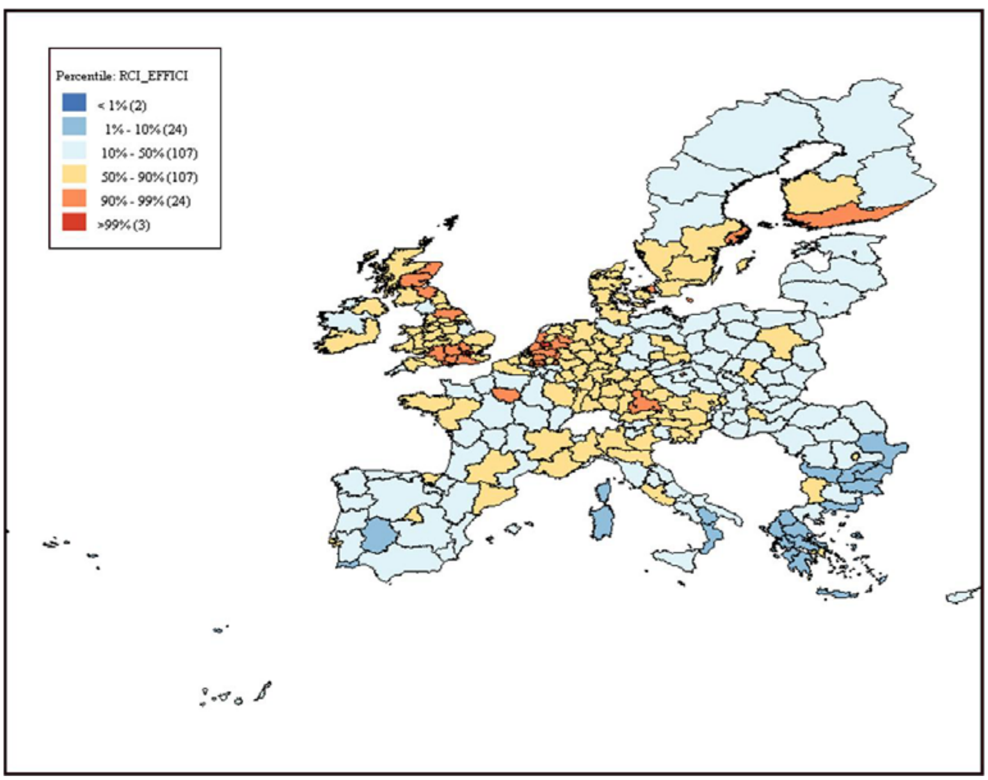


Figure 6: RCI-efficiency percentile map

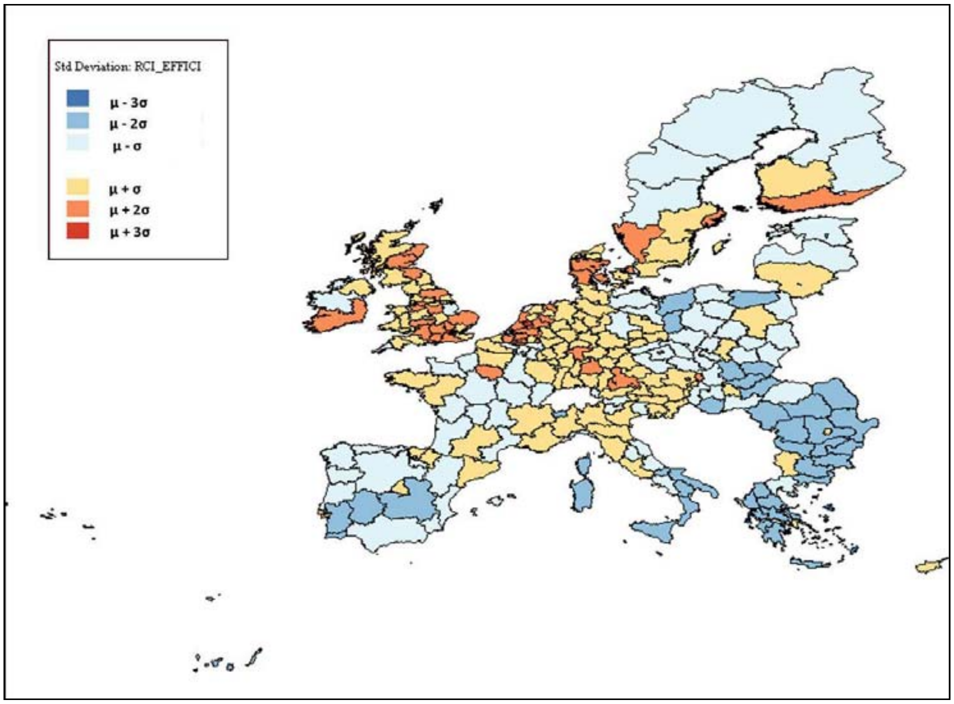


Figure 7: RCI-efficiency standard deviation map

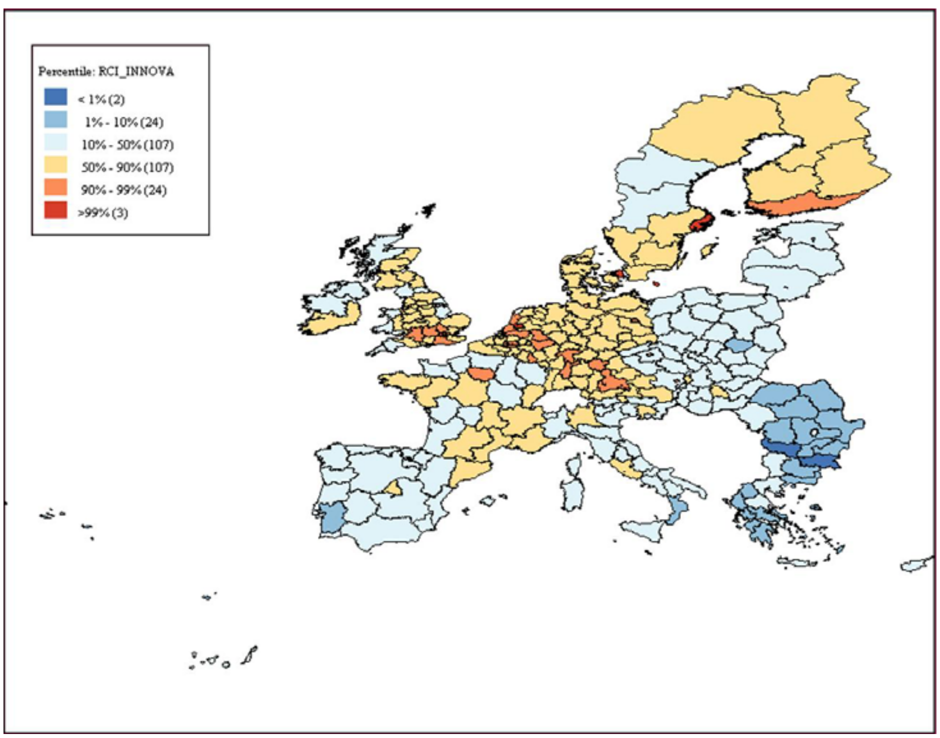


Figure 8: RCI-innovation percentile map

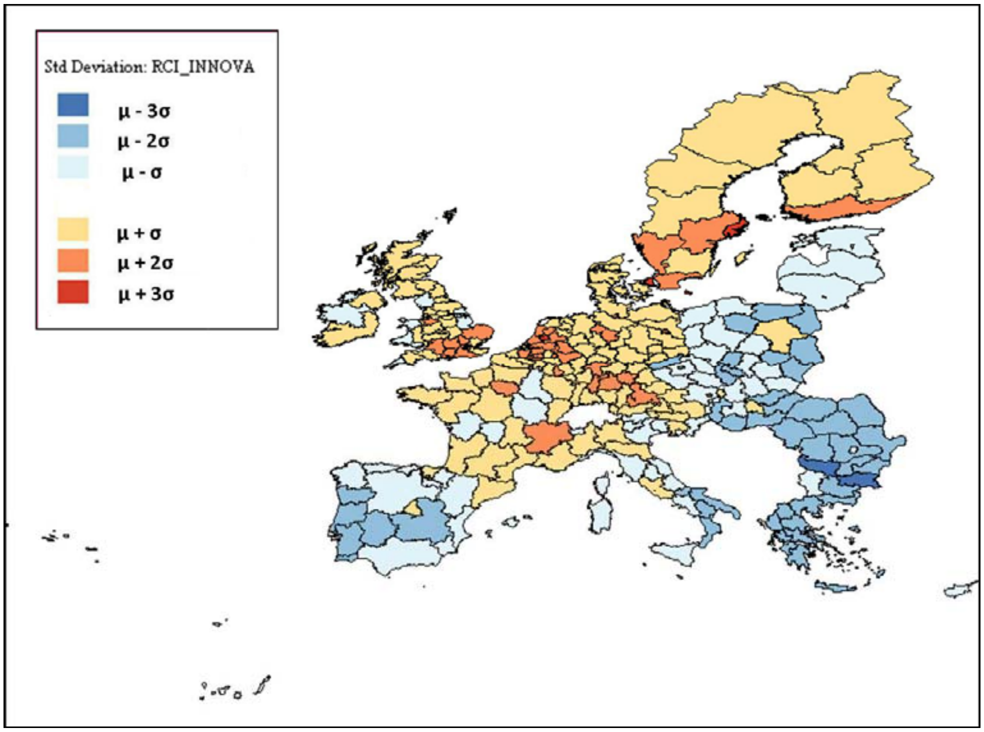


Figure 9: RCI-innovation standard deviation map

From this simple visualization exercise, we can conclude that there is indication for the existence of clusters of regions with similar performance. Two main areas can be detected: one which includes central and northern countries, the other including Bulgaria, Romania and Greece as well Central Eastern European regions. Italy, France and Spain show heterogeneous patterns quite different from other areas.

1.1.2 Assessment of spatial autocorrelation

Another stage of ESDA aims at identifying the structure of the spatial correlation that better describes the data. Spatial autocorrelation is concerned with exploring the existence of a systematic pattern in the spatial distribution of a variable. Positive spatial autocorrelation corresponds to neighboring areas having significant positive correlation and being more alike, while negative autocorrelation suggests the opposite. In the case of RCI, a proof for positive spatial autocorrelation could be taken as evidence of the existence of a spillover effect among groups of EU NUTS 2 regions.

When measuring spatial dependency in a dataset the first step is the definition of spatial relationships which exist between any set of points or areas (Haining, 2003). Next, there are several ways of measuring spatial dependence. These measures can be either applied to the whole area under analysis or to subsets of areas, if the entire area is highly heterogeneous.

Geometric/spatial criteria are used to define relationships between objects if physical proximity is expected to be the main driver of similarity. Spatial relationships can be simply represented in the form of a binary connectivity matrix **C,** which is a n x n symmetric matrix with n number of objects/areas. If two objects i and j are adjacent, they are in relation with each other and:

$$c(i, j) = c(j, i) = 1$$

with $c(i,j)$ denoting the cell (i,j) of matrix **C**. Otherwise $c(i,j) = 0$. Matrix **C** describes the so called first-order adjacencies as it indicates pairs of objects which are directly connected. Matrix $\mathbf{C}^2 = \mathbf{C} \times \mathbf{C}$ describes second-order adjacencies as it identifies all pairs of objects that can reach each other in two steps. Values in $\mathbf{C}^2$ can be integer numbers higher than 1; they are the number of different pathways from object i to object j and vice-versa. Higher-order adjacencies can be found similarly by matrices $\mathbf{C}^3$, $\mathbf{C}^4$, etc. More complex criteria than presence/absence of direct adjacencies may be used to define to define spatial dependency. They all make use of a weight matrix, **W**. Examples of weight matrix are[2] (Anselin, 1988; Getis and Aldstadt, 2002; Haining, 2003):

- Distance: $w(i, j) = 1/d_{i,j}^{\delta}$ where $d_{i,j}$ is a certain distance between objects/areas i and j and $\delta \geq 0$;
- Common border: $w(i, j) = \left(l_{i,j}/l_i \right)^{\tau}$ $\tau \geq 0$ where $l_{i,j}$ is the length of the common border between i and j and l_i is the length of border i;
- Combined border and distance weighting: $w(i, j) = \left(l_{i,j}/l_i \right)^{\tau} d_{i,j}^{-\delta}$

[2] If objects are areas, distances are computed between area centroids.

The calculation of a spatial autocorrelation measure is highly dependent upon the definition of neighbors. The weights matrix $w(i, j)$ could be a contiguity-based or distance-based spatial matrix. A contiguity-based spatial weight matrix implies a definition of a neighbor based on sharing a common boundary or being situated within a given distance band. Common contiguity types are a "rook" variety (only pure borders) or a "queen" variety (both borders and common vertices)[3]. Rook is a more stringent definition of polygon contiguity than queen—for rook the shared border must be of some length, whereas for queen the shared border can be as small as one vertex. A distance-based spatial weight matrix implies the definition of a distance; the neighborhood is then inversely related to the distance between points or between polygon centroids $1/d_{i,j}^{\delta}$.

Another possibility is the so called 'nearest neighbors' method where each point (region) is linked to its k ($k = 1, 2, 3, \ldots$) nearest neighbors. The k-nearest neighbors weight matrix **W** is computed by defining weights w_k^* which depend on the number of neighbors k (Ertur and Koch, 2006):

$$\begin{cases} w_k^*(i, j) = 0 & \text{if } i = j \\ w_k^*(i, j) = 1 & \text{if } d_{ij} \le d_i(k) \\ w_k^*(i, j) = 0 & \text{if } d_{ij} > d_i(k) \end{cases}$$

where $d_i(k)$ is a cut-off distance which depends on each point i and is the shortest distance between point i and its neighbors such as point i has exactly k neighbors in a circle centered in point i itself with radius $d_i(k)$. The k-nearest neighbor criterion ensures that each point has exactly the same number (k) of neighbors, but note that if point i is one of the k nearest neighbours of j, this does not imply that j is the k nearest neighbors of i (Haining, 2003). This method is also based on the choice of a particular type of distance. The distance adopted for the RCI case in the ESDA analysis is the Euclidean distance between population weighted centroids of the regions.

[3] These terms are derived from an analogy to a chess board, where the rook neighbors would be the four locations to the North, South, East and West, and the queen neighbors would also include the corner elements (for a total of eight neighbors).

In practice, it is nearly impossible to choose a "best" spatial proximity matrix and typically one assesses the sensitivity of the results to the selection of weights.

A number of statistics for overall clustering exists, where a null hypothesis is spatial randomness in the distribution of data values, such as the global Moran index or the Geary index (Haining, 2003).

1.1.3 Global measures

Moran's global index (Moran's I) is one of the oldest and most familiar indicators of spatial autocorrelation (Moran, 1950). It compares the value of the variable (in our case the RCI scores) at any one location with the value at all other locations. If neighboring units over the entire study area have similar values, then the statistics should indicate a strong positive spatial autocorrelation. If neighboring units over the entire study area have dissimilar values, then the statistics should indicate a strong negative spatial autocorrelation. This statistic is essentially a cross-product correlation measure that incorporates "space" by means of a spatial weights matrix.

Moran's I is defined as:

$$I = n \frac{\sum_{i,j} w(i,j)\left(X_i - \bar{X}\right)\left(X_j - \bar{X}\right)}{\sum_i \left(X_i - \bar{X}\right)^2} \qquad \sum_{i,j} w(i,j) = 1 \qquad (1)$$

where n is the number of spatial units (NUTS2 regions in our case) indexed by i and j, X is the variable of interest, $\bar{X}$ is the average value of X across all the points in the area and $w(i,j)$ is the cell (i,j) of the matrix **W** of spatial weights.

Similar to correlation coefficients, Moran's index varies between -1 (perfect dispersion) and +1 (perfect spatial correlation). Perfect dispersion means that high values are always surrounded by low values and vice-versa. Perfect correlation means that there's always a concentration of high-high or low-low. Note indeed that the numerator of (1) is positive when X_i and X_j are both greater that or less than the mean value of X. This means that the index I does not distinguish between the concentration of high values and the concentration of low values. When the analyst is interested in this distinction the Getis-Ord statistic is preferable (Haining, 2003).

Under the null hypothesis of no spatial correlation, the expected value of the Moran's I depends only on the number of objects n, $\mathrm{E}[I]=-\frac{1}{n-1}$. In the RCI case, the number of NUTS2 regions is n=267 (the 4 French overseas regions have been excluded from the analysis) so $\mathrm{E}[I]=-0.0038$. Values of I larger than the expected value indicate positive spatial autocorrelation, while values smaller than the expected indicate negative spatial autocorrelation.

Inference on I is based on the permutation approach, assuming that, under the null hypothesis, each observed value could have occurred at all locations with equal likelihood. A reference distribution is empirically generated for I, from which the mean and standard deviation are computed. In practice this is carried out by permuting the observed values over all locations and by re-computing I for each new sample (Anselin, 1995). Statistically significant values of the spatial autocorrelation indexes are evidence of spatial dependency and the potential existence of spillover effects.

Moran's I statistics are shown in Table 1 for the RCI and its three sub-indexes. Moran's I statistics have been obtained using different spatial matrices in order to test for the influence of different weight matrices on results: the queen and rook contiguity matrices, the matrix of Euclidean distance between centroids and the k nearest neighbors method. The queen and rook contiguity matrices have been used up to the third-order adjacency and the k nearest neighbors method was carried out with k = 5, 10, 15, 20. Table 1 shows results only for the first-order adjacency for the contiguity matrices and for k=5 for the nearest neighbors method. As evidenced by Le Gallo and Ertur (2003) European regions have on average about 5 contiguous neighbors so that the choice of k=5 yields a ring around each region of approximately the first order contiguous regions.

It appears that the RCI is positively spatially autocorrelated as all statistics are significant with p=0.0001. A total number of 9999 permutations have been used for all cases.

Similar results, all associated to significant p-values, have been obtained with higher order contiguities and for different number of nearest neighbors[4].

Table 1: Moran's *I* statistic for RCI and RCI sub-indexes.
Different spatial weights matrices are used for comparison.

	Type of spatial weight matrix	Moran's I	mean	standard deviation	p-values
RCI	queen contiguity matrix	0.663	-0.0044	0.0427	0.0001
	rook contiguity matrix	0.6629	-0.0037	0.0431	0.0001
	k-nearest neighbors (5)	0.7451	-0.0044	0.0361	0.0001
RCI basic sub-index	queen contiguity matrix	0.82	-0.0043	0.0428	0.0001
	rook contiguity matrix	0.824	-0.0034	0.0429	0.0001
	k-nearest neighbors (5)	0.8595	-0.0039	0.0358	0.0001
RCI efficiency sub-index	queen contiguity matrix	0.5272	-0.0041	0.043	0.0001
	rook contiguity matrix	0.5278	-0.0037	0.043	0.0001
	k-nearest neighbors (5)	0.6332	-0.0045	0.0362	0.0001
RCI innovation sub-index	queen contiguity matrix	0.6726	-0.0038	0.0042	0.0001
	rook contiguity matrix	0.6732	-0.0041	0.0432	0.0001
	k-nearest neighbors (5)	0.7406	-0.0042	0.0358	0.0001

Global indicators of spatial autocorrelation as the Moran's index give a unique measure of spatial association for the whole dataset, useful for the characterization of the study area as a whole. They do not give information as to where the clusters or outliers are located or the type of spatial correlation that is most important (e.g. correlation between high or between low values) (Anselin et al., 2007). One can wonder which regions contribute more to the global spatial autocorrelation, if there are local spatial clusters of high or low values, and to what point the global evaluation of spatial autocorrelation hides atypical localizations, i.e. regions or groups of contiguous regions deviating from the global pattern of positive spatial autocorrelation. Further, if a high number of regions

[4] Complete results are available from the authors.

is examined, as in the RCI case, there exists a higher probability for the existence of different sorts of spatial autocorrelation in different sub-regions. In this case, local indicators of spatial autocorrelation and the Moran scatter plot can be used.

1.1.4 Moran's scatterplots

Moran's *I* provides a global measure of spatial correlation without any associated inference and is intrinsically based on the assumption of stationarity or structural stability of the process over the space. The dependence structure is said to be stationary if the nature of similarity between nearby values of the variable of interest is independent on where values are measured (Haining, 2003).

Being global, the Moran's *I* can be correctly interpreted only if the assumption of stationarity is realistic and Moran's measure cannot be used to assess the local structure of spatial autocorrelation. Instead, it may be interesting to detect local spatial clusters of high or low value, which areas contribute more to the global spatial autocorrelation and to what extent the global measure masks atypical groups of regions or, as recently nicely defined by Ertur and Koch (2006), 'pockets of local nonstationarity'.

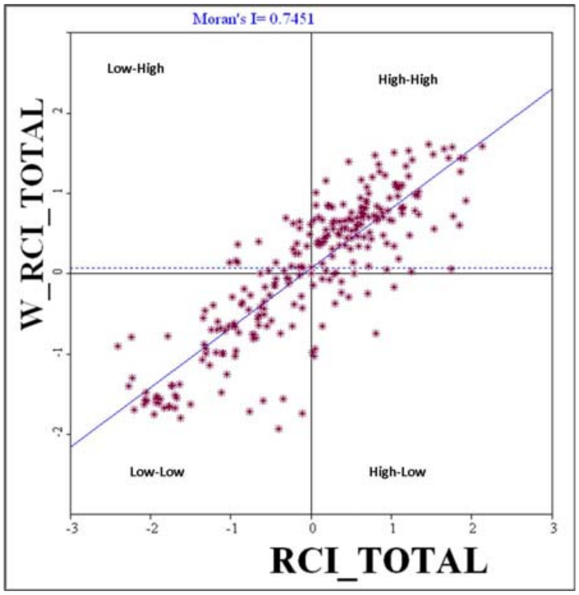
A simple way to study local spatial instability is by means of the Moran scatter plot (Anselin 1995, Ertur and Koch, 2006) which visualizes the slope in a scatter plot of the spatially lagged variable on the original one. The rationale is to compare (normalized) values of the variable in an area (region) with the average of its neighbors, constructing a bidimensional plot of *z* (normalized values) by W_z (average of neighbors), divided into four quadrants. The four different quadrants correspond to four types of local spatial association between one region and its neighbors (defined by a certain weight matrix): high-high, low-low, low-high or high-low[5]. Atypical localizations are those which fall in the high-low or low-high quadrants.

In this kind of plots the spatial lag of the variable (W_z) is shown on the vertical axis and the original variable (z) on the horizontal axis. The 'spatial lag' refers to the values of a location's neighbors. If constructed based on normalized values, it allows for an analysis of the behavior of the spatial variability.

[5] High and low mean respectively above and below the average value.

Low-High regions:

BE34	PL42
CZ02	PL43
CZ03	PL63
CZ04	UKD1
CZ06	UKF3
EE00	UKK3
ES24	UKM6
FR21	
FR25	
FR26	
FR43	
FR61	
FR63	
FR72	
FR81	
HU22	
ITC1	
ITC2	
ITD1	
ITD2	



High-Low regions:
AT22
CZ01
DE41
ES21
ES30
ES51
FR61
FR62
FR71
FR82
HU10
ITC4
ITD3
ITD5
ITE4
PL12
PT17

Figure 10: RCI - Moran's scatter plot
(Results are based on 5-nearest neighbors)

The interpretation of the quadrants is as follows:

- Q1 (positive values, positive neighbors' mean) and Q3 (negative values, negative neighbors' means) – indicate points of positive spatial association, i.e. neighbors have similar values, either high-high (Q1) or low-low (Q3);

- Q2 (positive values, negative neighbors' means) and Q4 (negative values, positive neighbors' means) – indicate points of negative spatial association, i.e. a region has neighbors with distinct, different values, either low-high (Q2) or high-low (Q4).

In the Moran's scatterplot the level of global spatial autocorrelation is visualized as well. As can be seen from eq. (1), Moran's *I* is indeed formally equivalent to the slope coefficient of the linear regression of the spatial lag of the variable (W_z) and z so that the higher the slope, the stronger the global spatial autocorrelation (Ertur and Koch, 2006).

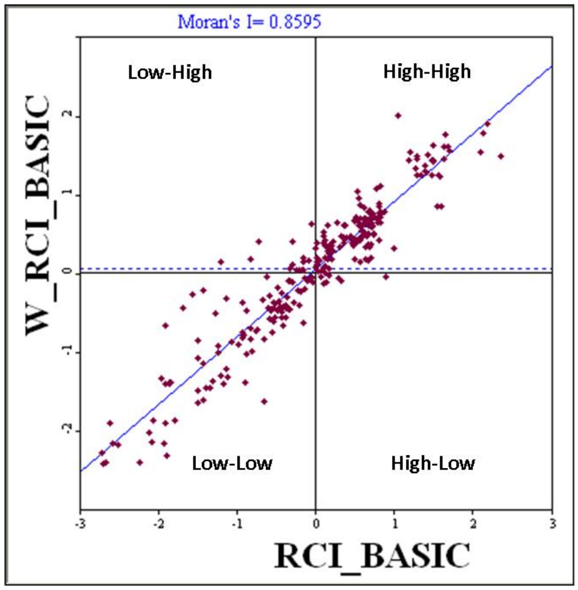
Figure 10 shows the Moran scatterplot for the total RCI computed using the 5 nearest neighbor matrix as weight matrix. Most regions are located in quadrants Q1 and Q3, confirming the evidence of positive spatial autocorrelation. Atypical regions, i.e. regions that deviate from the global spatial association pattern and belong to the quadrant low-high or high-low, are about 16% of all regions (17 regions in the high-low quadrant and 27 regions in the low-high).

The group of high-low regions includes, among others, 4 Italian regions (ITC4, ITD3, ITD5, ITE4), 4 French regions (FR61, FR62, FR71, FR82), 3 Spanish regions (ES21, ES30, ES51) and some capital regions - the Czech Republic (CZ01), Portugal (PT17) and Hungary (HU10). Regions in the low-high quadrant are 8 French regions (FR21, FR25, FR26, FR43, FR61, FR63, FR72, FR81), 1 Spanish region (ES24), 4 Italian regions (ITC1, ITC2, ITD1, ITD2), 4 UK regions (UKD1, UKF3, UKK3, UKM6), 4 Czech regions (CZ02, CZ03, CZ04, CZ06) at the border with Germany, 3 Polish regions (PL42, PL43, PL63), 1 Belgium region (BE34) at the southern border with Luxembourg and France, Estonia (EE00) and one region in Hungary (HU22) at the western border with Austria.

These findings confirm the remarks made in Section 1.1.1 based on the visualization of the maps: Italy, France and Spain show heterogeneous performances and lack of indication for the presence of regional spillovers. Not surprisingly capital regions, especially in new Member States such as the Czech Republic and Hungary are surrounded by lower performing regions. The concentration of economic activity in those regions is very strong and spillover effects to the neighboring regions are not yet evident. In any case, positive spatial associations – either high-high or low-low - are prevailing across the major part of EU NUTS2 regions.

Moran scatterplots for the three separate sub-indices of RCI are shown in Figure 11. The picture with highest positive spatial autocorrelation is the one referring to the basic sub-index (top quadrant of Figure 11). As aforementioned, this is due to the nature itself of the first group of pillars which is mostly measured at the country level resulting in an index which is, by construction, the most spatially autocorrelated. For the other two sub-indices the scatterplots indicate a slightly higher spillover effect for the innovation sub-index than for the efficiency sub-index: a higher number of regions fall in the quadrants Q1 and Q3 for the innovation sub-index (87% of all regions) than for the efficiency one (73% of all regions).

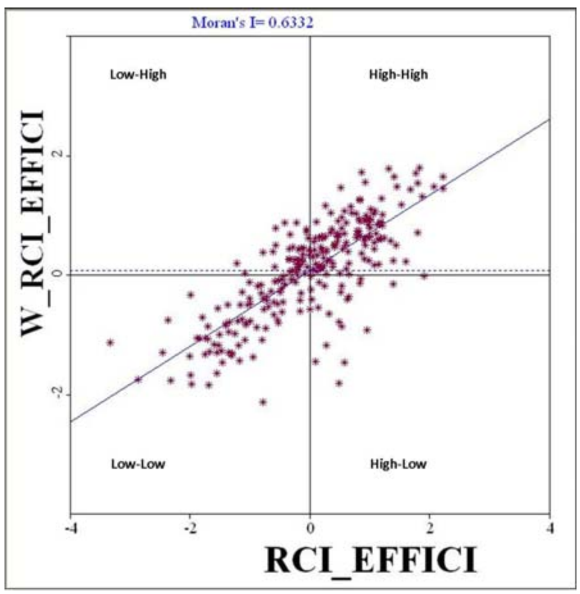
Low-High regions:
CZ02
CZ03
CZ04
CZ06
EE00
ES24
FR25
FR52
FR53
FR63
HU22
PL42
PL43
SK01
UKF3
UKK3
UKM5
UM6



High-Low regions:
AT13
CZ01
ES30
ES42
ES51
FR71
FR82
UKM2

Low-High regions:

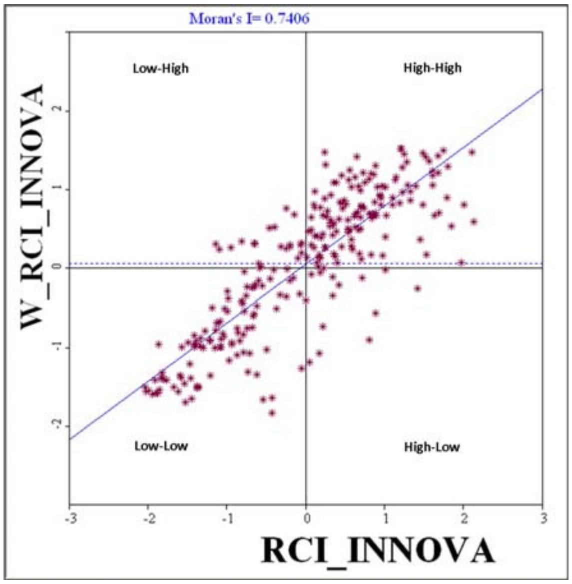
AT34	FR81
BE32	ITC2
BE34	ITC3
BE35	ITD1
IE01	ITD4
CZ02	ITE2
CZ03	ITE3
CZ04	PL43
CZ06	SE31
DE80	SE32
DED1	SE33
DEE0	SK02
EE00	UKF3
ES24	UKK3
FI13	
FI1A	
FI20	
FR21	
FR23	
FR25	
FR61	



High-Low regions:

BG41	PL12
CY00	PL22
CZ01	PT17
DE30	RO32
DED1	
ES21	
ES22	
ES30	
ES51	
FR10	
FR62	
FR71	
FR82	
GR30	
HU10	
ITC4	
ITE1	
ITE4	
LU00	

Low-High regions:
AT11
AT12
CZ04
CZ06
EE00
ES24
FR21
FR26
FR53
FR63
HU22
IE01
ITC2
PL42
PL43
UKD1
UKE1
UKF3
UKK3
UKL1



High-Low regions:
AT13
AT22
CZ01
ES21
ES30
ES51
FR52
FR71
HU10
ITC4
ITD5
ITE4
PL12
SI02
SK01

Figure 11: Moran scatterplots for the three RCI sub-indices
(Results are based on 5-nearest neighbors)

1.1.5 LISA measures

The Moran's I scatter plot is the first step in examining different types of spatial association. Despite the information it provides, it does not give any indication about statistical significance of spatial autocorrelation. Moran's I scatter plots are simply descriptive, if one region is surrounded by lower or higher values it is assigned to the 2nd or 4th quadrant whatever the difference between the values.

Local Indicators of Spatial Association – LISA, proposed by Anselin in 1995, are local measures providing statistical inference. LISA are indicators of spatial dependence associated to different localizations of the spatially distributed variables. They allow for the decomposition of global indicators, such as Moran's I, into the contribution of each individual observation and assess the significance of the local clusters (high-high or low-low) or local spatial outliers (high-low or low-high) as identified also in the Moran's scatter plot in a descriptive way. Significance is based on a conditional permutation approach. In practice LISA decompose global indicators into a sum of local patterns. If the process is stationary, the local spatial patterns will be in line with the global index. The presence of areas which strongly deviate from the average patterns are detected as outliers as they indicate locations that contribute to the global index more than their expected share. While the global tests suggest if there is spatial correlation, the local tests can show where this is occurring (Perry et al., 2006). For each point a local indicator is computed and their sum, across all points, is proportional to a global indicator of spatial association. This means that the individual components of LISA are related to the global statistic of spatial association. It is exactly for this reason that LISA outliers can be associated with those regions which mostly influence the global measure, such as the Moran's *I*.

Moran's local index can be expressed for each point *i* (region in our case) as:

$$I_i = \frac{\left(X_i - \bar{X}\right)\sum_{j=1}^{n} w(i,j)\left(X_j - \bar{X}\right)}{\sum_{j=1}^{n}\left(X_j - \bar{X}\right)^2} \tag{2}$$

where the summation over *j* is such that only neighboring values of point *i* are included (the definition of neighbors depends on the definition of contiguity as for the global Moran's *I*). Similarly to the previous analysis, 5 nearest neighbors matrix is used as weight matrix.

Under the null hypothesis of no spatial association the expected value and the variance of I_i can be analytically derived. The expected value at each point *i* is: $\mathrm{E}[I_i] = {-\sum_j w(i,j)} \big/ {(n-1)}$

A test of significance of local spatial association is based on the conditional permutation approach (Ertur and Koch, 2006). The value X_i at a certain point *i* is held fixed and the remaining values are randomly permuted over all locations. In the RCI case 9999 permutations are used to compute the empirical distribution function and, from that, to compute *p*-values.

LISA cluster maps show regions with significant local Moran statistics, classified in the four groups of spatial correlation (high-high, low-low, high-low and low-high). Figure 12 shows the LISA Cluster map for the RCI. We can find in color the local clusters or outliers which are 'significant' (at the level $\alpha = 0.05$), based on the Local Moran statistic.

Local clusters of low values are located in Bulgaria, Romania, Greece and Cyprus, south-eastern regions of Hungary and Slovakia, southern part of Italy, Portugal and Spain. These areas show significant low-low clusters, meaning that regions with low RCI scores are surrounded by low scoring regions. On the other hand, local clusters of high RCI values are found in regions in the Netherlands, parts of Germany, Belgium, Denmark, the southern part of United Kingdom, Finland and Sweden.

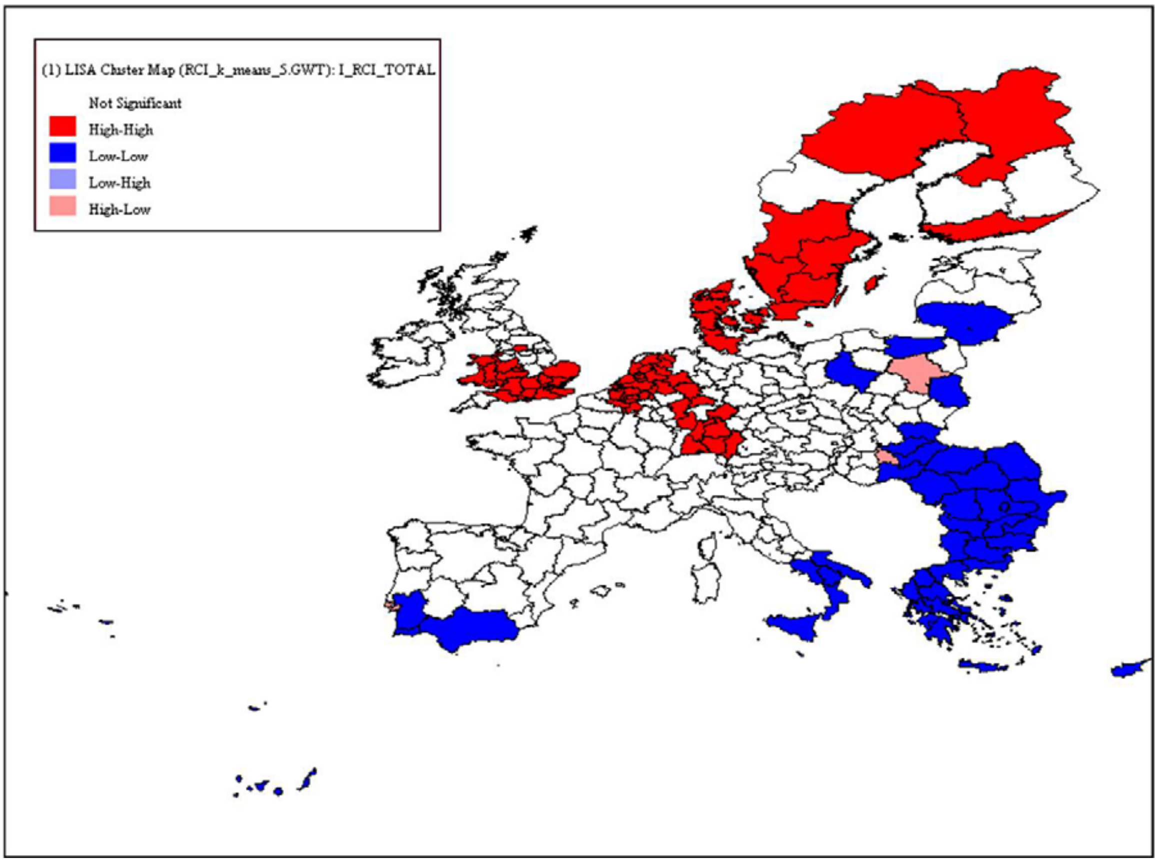
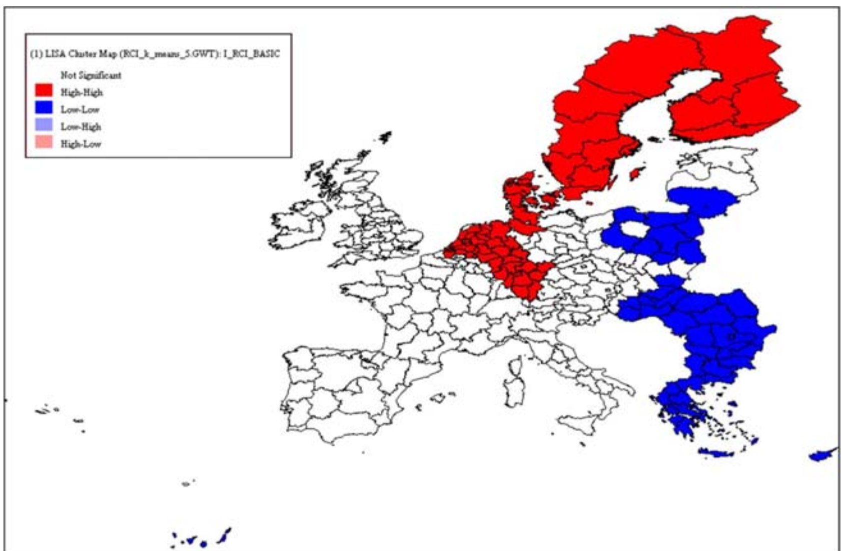
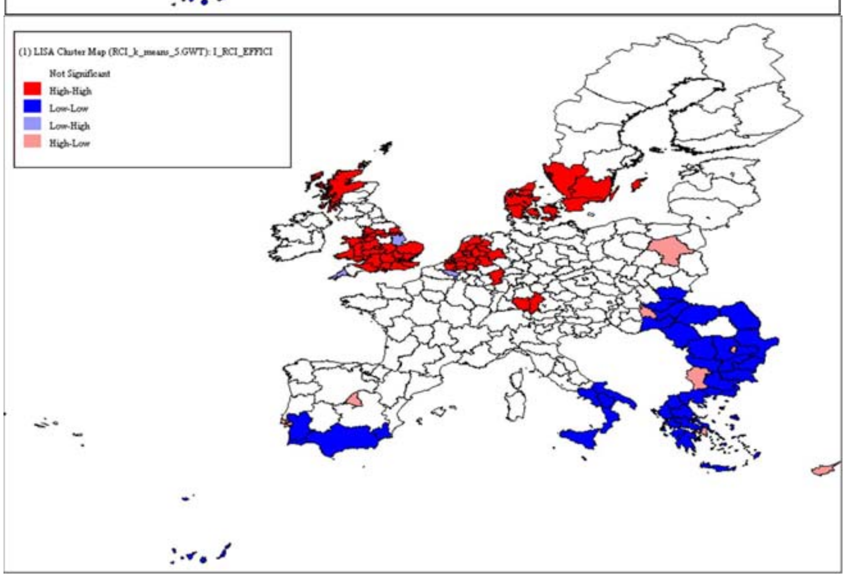


Figure 12: LISA Cluster Map for RCI total
(Results are based on 5-nearest neighbors)

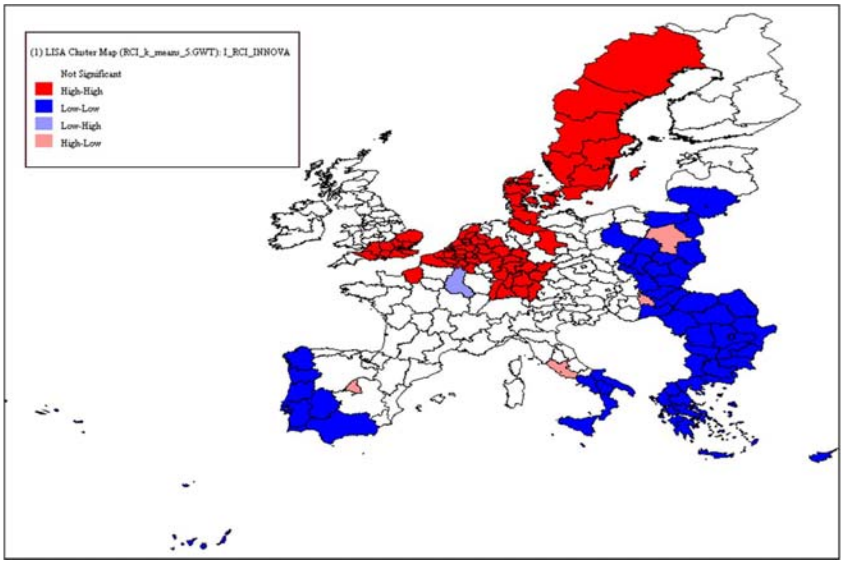
Three separate LISA analyses for the RCI sub-indices are provided in Figure 13. The large regions of high-high and low-low autocorrelations of the basic RCI sub-index (larger than the ones in the other two sub-indexes) are intrinsically due to the nature of the sub-index, where three out of five dimensions are measured at the country level, as aforementioned. It is interesting to note that in the efficiency and innovation LISA maps some region capitals show up as high-low areas: this happens for Spain (Madrid region ES30), Italy (Rome region ITE4), Greece (Athens region GR30), Bulgaria (Sophia region BG41), Romania (Bucharest region RO32), Hungary (Budapest region HU10) and Poland (Warsaw region PL12).



RCI basic



RCI efficiency



RCI innovation

Figure 13: LISA Cluster maps for the three RCI sub-indices
(Results are based on 5-nearest neighbors)

Table in Appendix A shows LISA statistics for RCI total for all the European regions with their associated significance level.

1.2 The variogram

The variogram is a function which describes spatial dependencies of a set of georeferenced values (Cressie, 1984; Thompson, 1992). It has been developed within the geological sciences where spatial variation has traditionally been summarized using the 'variogram' in place of the covariance function. The variogram and its estimation is an essential ingredient of the spatial prediction model known as 'Kriging' as it is the necessary step towards the prediction of values over space (or time). In our case we are interested in the variogram itself rather than in the prediction of values in unobserved locations. The variogram provides in fact a global description of spatial dependency and adds information to other global measures of spatial correlation.

The variogram is defined as the variance of the difference of the value of the variable of interest *y* at separate points in the area of interest:

$$2\gamma(h) = \mathrm{Var}[y_{i+h} - y_i] \qquad (3)$$

where y_i is the value of y at point i and y_{i+h} is the value of y in a point which is separated from i by a distance h [6]. The function $\gamma(h)$ is called semi-variogram and describes the spatial dependency structure. The structure is said to be 'second-order stationary' if the expected value of y is constant over the area under analysis - E[y_i] = μ for all the points i - and the covariance of the y-values at any two points which are separated by h depends only on the distance h,. With this assumption the semi-variogram is considered to be valid over the entire set of data and the relationship between the semi-variogram and the covariance of y is:

$$\begin{aligned} &\mathrm{Cov}[y_{i+h}, y_i] = \mathrm{E}[y_{i+h} \cdot y_i] - \mathrm{E}^2[y_i] = \mathrm{E}[y_{i+h} \cdot y_i] - \mu^2 \equiv C(h) \\ &\gamma(h) = \mathrm{Var}(y) - C(h) \end{aligned} \tag{4}$$

A simple method to estimate the (semi-)variogram is:

$$2\hat{\gamma}(h) = \frac{1}{n(h)} \sum_i (y_{i+h} - y_i)^2 \tag{5}$$

where the summation is over all distinct pairs of points that are h distance apart and $n(h)$ is the number of pairs that distance apart. Variogram values will be small the more alike values separated by distance h are (Haining, 2003). The variogram increases as the values get more and more dissimilar so that $\hat{\gamma}(h)$ tends to increase as h increases in the presence of a certain level of spatial autocorrelation of the variable under analysis. $\hat{\gamma}(h)$ is then a (estimated) value of dissimilarity as can also be seen from the opposite relation between the variogram and the covariance function $C(h)$ in (4).

The semi-variogram function is generally estimated by fitting the best curve for the points $\{h, \hat{\gamma}(h)\}$ with a function which has to be positive definite. The plot of $\hat{\gamma}(h)$ as a function of h provides a graphical description of the dependence structure in the data for different distances. Figure 14 shows an example of a semi-variogram (solid line) estimated from a set of points $\{h, \hat{\gamma}(h)\}$ (black diamonds in the Figure). In this case the spatial dependence is strong at short distances and rapidly decreases as h increases up to a certain distance – the range - beyond which the covariance levels off to nearly zero and

[6] In the present study we are assuming isotropy in the spatial variation.

the semi-variogram reaches its asymptotic value which is the variance of y over the area - Var(y). This value is called the 'sill' of the variogram.

In general, the shape of the (semi-)variogram is informative on: i) the speed at which autocorrelation decreases as the distance increases, from the slope of the semi-variogram, and ii) the maximum distance beyond which correlation can be considered null, from the value of the range. These are two important pieces of information that complement classical ESDA measures such as Moran's or LISA statistics.

Various analytical forms of variograms are generally used. Below is a list of the most common ones (de Marsily, 1986):

a) Power function: $h^b \quad 0 < b < 2$;

b) Spherical: $\begin{cases} \mathrm{Var}(y)\left[\frac{3}{2}\left(\frac{h}{r}\right) - \frac{1}{2}\left(\frac{h}{r}\right)^3\right] & 0 \le h \le r \\ \mathrm{Var}(y) & h > r; \end{cases}$

c) Exponential: $\mathrm{Var}(y)[1 - \exp(-h/r)]$;

d) Gaussian: $\mathrm{Var}(y)\{1 - \exp[-(h/r)^2]\}$;

where r is the range and Var(y) is the sill of the semi-variogram.

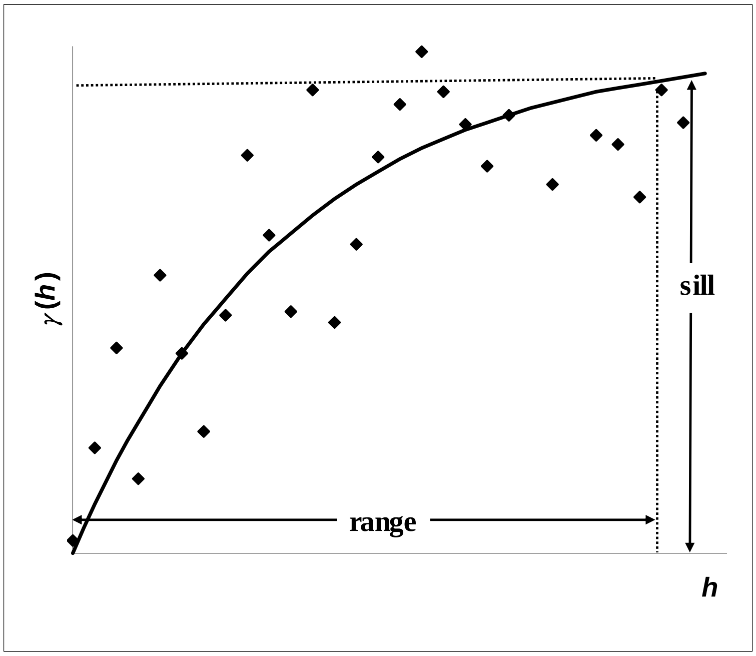


Figure 14: Model for a semi-variogram $\gamma(h)$

In the following different variogram analyses are provided for RCI. The structure of spatial dependence is explored separately for two sub-areas, as detected by the ESDA analysis, and three different types of distances between region centroids: the classical Euclidean distance, the distance along the real road or ferry network and the travel time distance. In all the analyses spatial correlation is assumed isotropic[7].

In the following analyses estimated variograms derive form a simple qualitative analysis; accordingly the estimated ranges have to be interpreted as a rough indication of maximum extent of spatial autocorrelation.

1.2.1 Euclidean distances

For the computation of the semi-variogram in the RCI case, we have selected two main sub-areas, mainly driven by considerations from the ESDA outcomes (Section 1.1.5). Within these sub-areas we assume that the second-order stationarity condition holds. The first sub-area – group A- comprises regions which have been classified as high-high by LISA analysis and includes eleven countries: AT-BE-DE-DK-FI-IE-LU-NL-SE-SI-UK.

[7] For the variogram analysis the following matlab scripts have been used: 1. 'variogram.m' (Copyright 2009, Wolfgang Schwanghart; Copyright 2006, The MathWorks) for calculating the isotropic and

The second sub-area – group B- comprises regions classified as low-low and includes ten countries: BG-CZ-EE-GR-HU-LT-LV-PL-RO-SK. The analysis of sub-area B is meant to further investigate the possible presence of ‘negative’ spillover effects where low performing regions negatively affect their neighbors.

In this step of the analysis distances between regions are computed as Euclidean distances between regions centroids.

Table 2 shows the basic settings for the estimation of the semi-variogram for the two groups of countries. The maximum distance at which the variogram is computed is set to 1/4 the maximum distance between regions centroids. The semi-variogram is estimated for four variables of interest: the RCI total score and the three sub-indices. The variance of the RCI scores is also shown in Table 2: in case of a structured spatial autocorrelation, they are the asymptotic values of the semi-variogram.

---

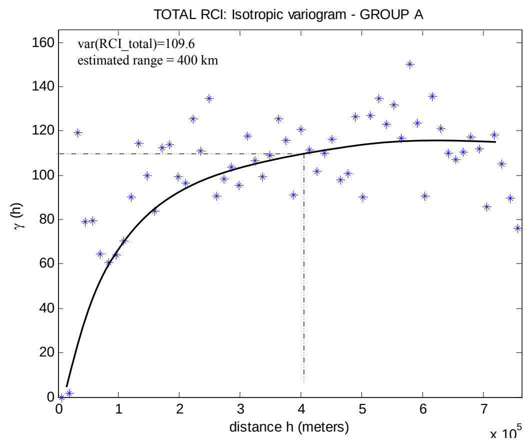
anisotropic experimental (semi-) variogram and 2. ‘ipdm.m’ (Copyright 2009, John D'Errico) for the computation of Inter-Point Distance Matrix.

Table 2: Basic settings for the semi-variogram analysis and estimated ranges with Euclidean distances between centroids

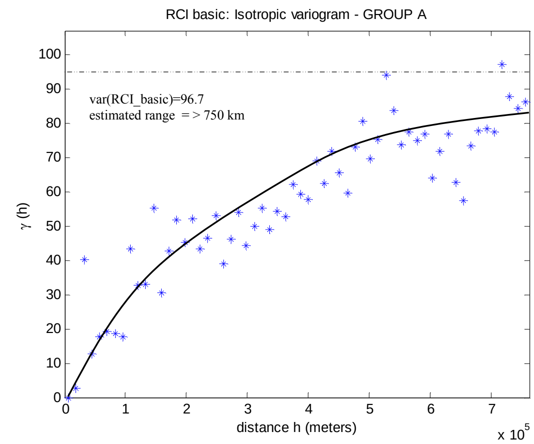
	MAXIMUM DISTANCE BETWEEN CENTROIDS*	MAXIMUM DISTANCE FOR VARIOGRAM ESTIMATION	variable of interest (min-max normalized values)	variance (asymptothic value for the semi-variogram)	estimated range (km)
GROUP A	3000 km	3000/4 =750 km	RCI_total	109.6	400
			RCI_basic	96.7	> 750
			RCI_efficiency	91.5	350
			RCI_innovation	190.5	300
GROUP B	2900 km	2900/4 = 725 km	RCI_total	251.7	na
			RCI_basic	223.3	700
			RCI_efficiency	178.7	na
			RCI_innovation	228.9	na
* Maximum distance computed as the Euclidean distance between two points in the area under investigation with minimum and maximum X-Y coordinates					

Results for regions in group A are shown in Figure 15 where estimated variograms are shown as solid black lines.

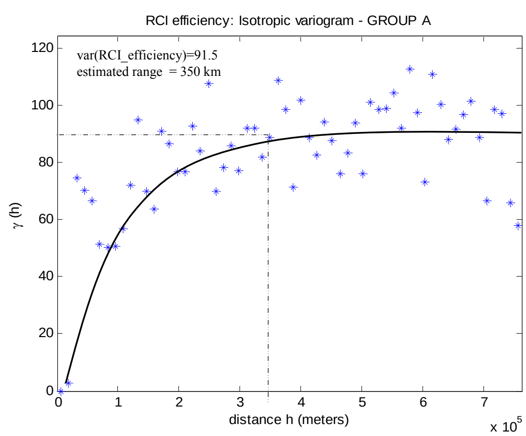
Figure 15 indicates the presence of a structure of spatial autocorrelation with an estimated range, that is the maximum distance beyond which the correlation is zero, always higher than 300 km. The highest range – > 750 km - is the one for the basic RCI sub-index, but this is due to the nature itself of the index that is mostly composed by indicators at the country level, as already remarked. The comparison between the efficiency semi-variogram and the innovation-one is very interesting. Both graphs (Figure 15 c and d) show a correlation which decreases as distances increase, which may be evidence of spill-over, but for the innovation sub-index the curve is steeper and the range slightly lower. This means that spatial correlation for the innovation sub-index decreases more rapidly than that for the efficiency sub-index. The shape of the semi-variogram for the total RCI score is very similar to that of the efficiency sub-index, with the same range and almost the same sill (asymptotic value).



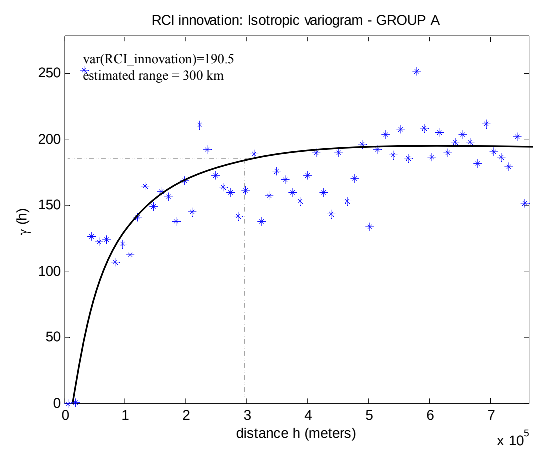
a) RCI total



b) RCI_basic



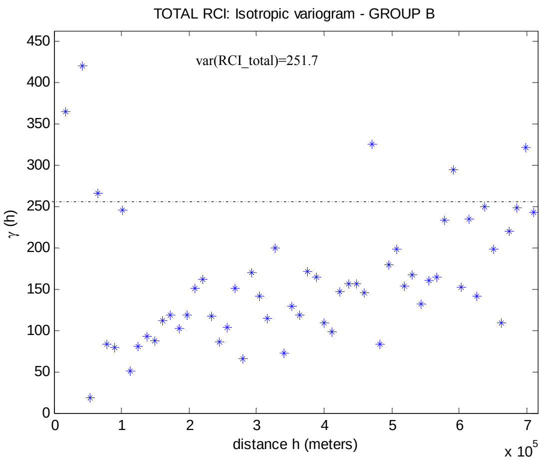
c) RCI_efficiency



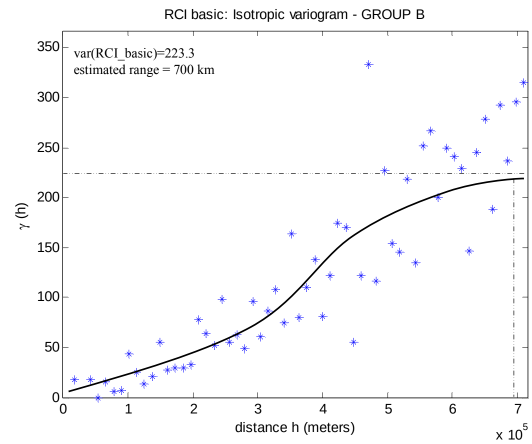
d) RCI_innovation

Figure 15: Semi-variogram estimation for regions in group A – Euclidean distances

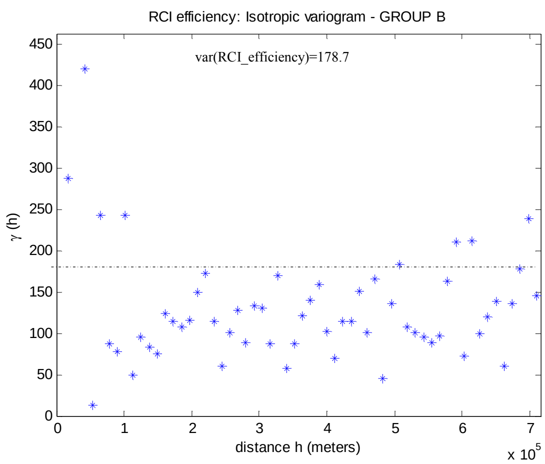
Results for regions in group B are shown in Figure 16. In this case no particular pattern can be picked out apart from the basic semi-variogram (Figure 16-b), but this is again related to the ‘less regional’ scale of measurement of the basic sub-index.



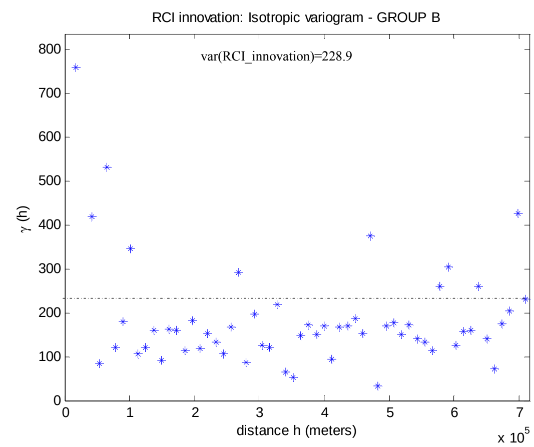
a) RCI total



b) RCI_basic



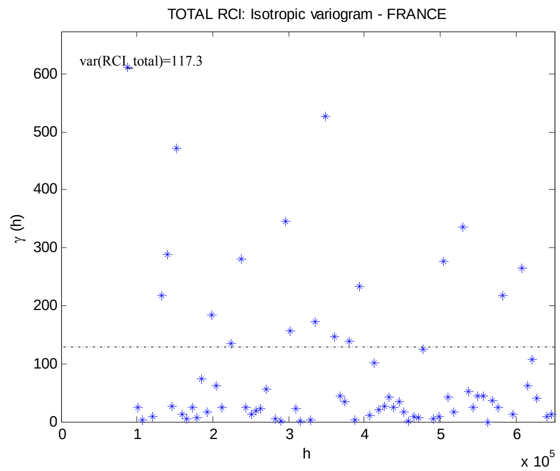
c) RCI_efficiency



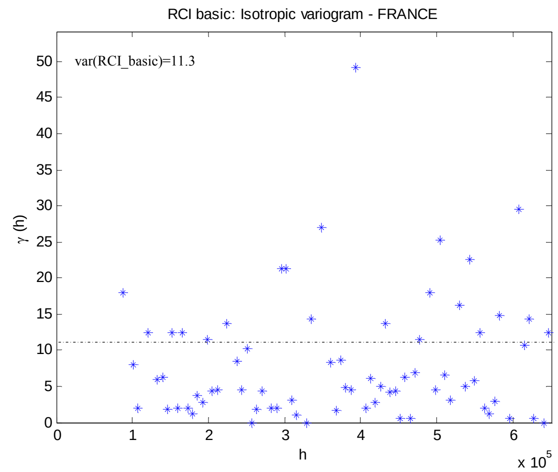
d) RCI_innovation

Figure 16: Semi-variogram estimation for regions in group B – Euclidean distances

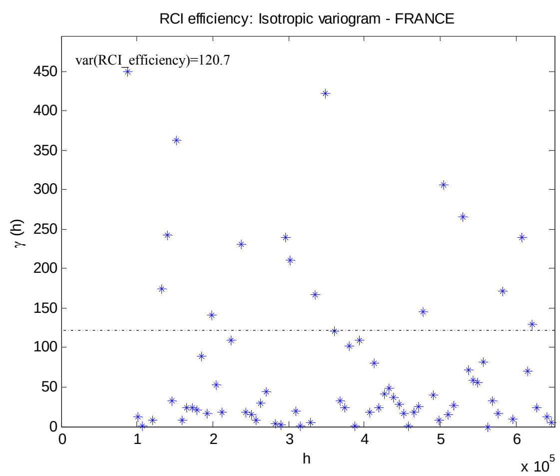
Similar analysis is carried out separately for France (Figure 17), Italy (Figure 18) and Spain together with Portugal (Figure 19) with non informative results as in the case of regions in group B.



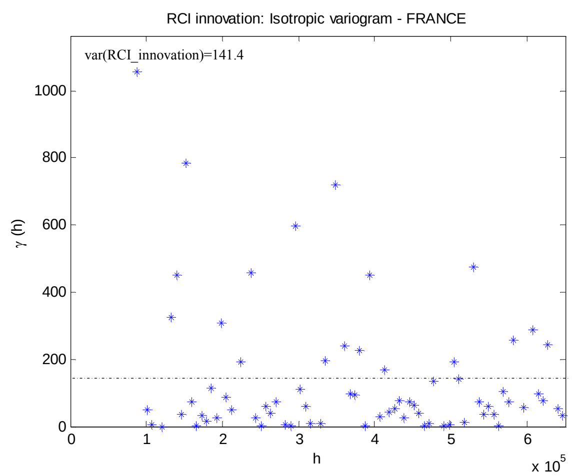
a) RCI total



b) RCI_basic

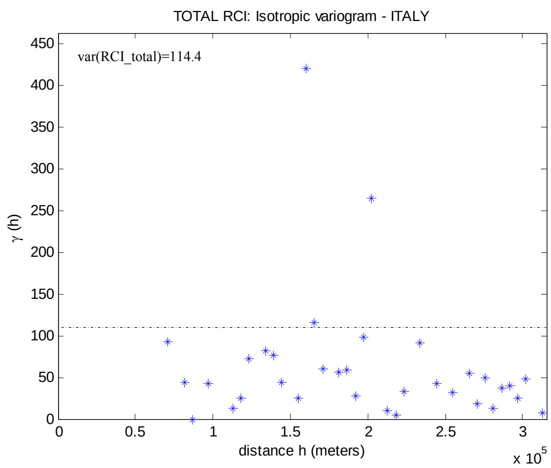


c) RCI_efficiency

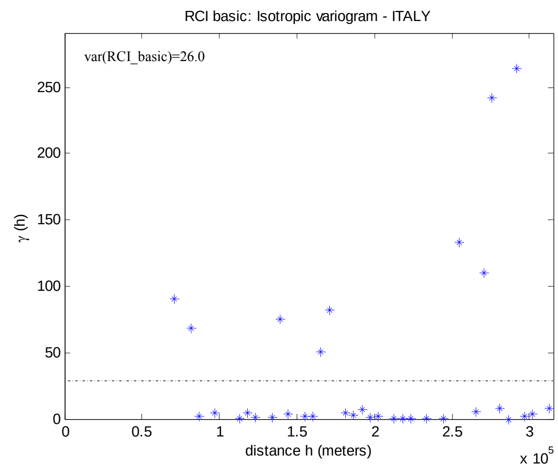


d) RCI_innovation

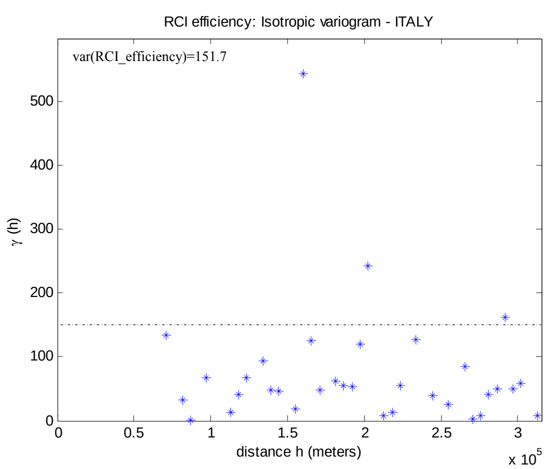
Figure 17: Semi-variogram estimation for regions in France – Euclidean distances (distances h in meters)



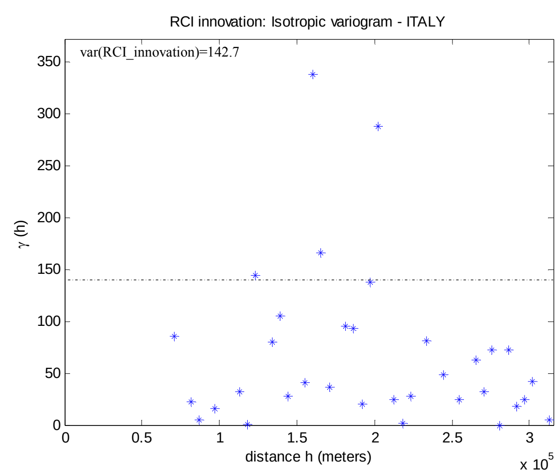
a) RCI total



b) RCI_basic

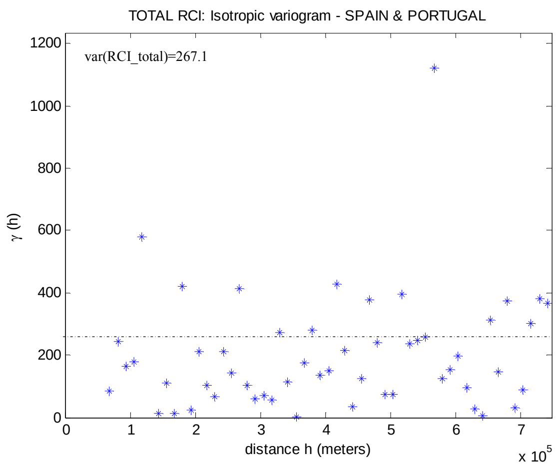


c) RCI_efficiency

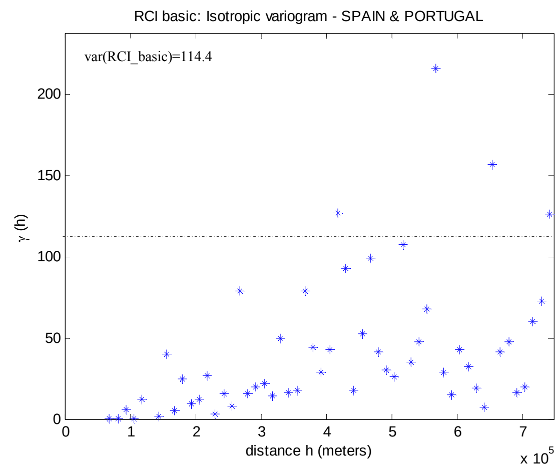


d) RCI_innovation

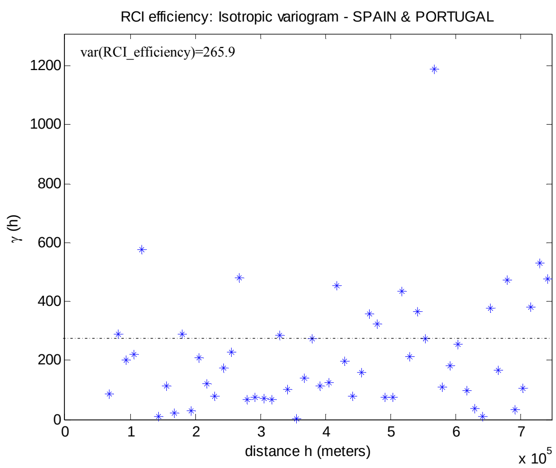
Figure 18: Semi-variogram estimation for regions in Italy – Euclidean distances



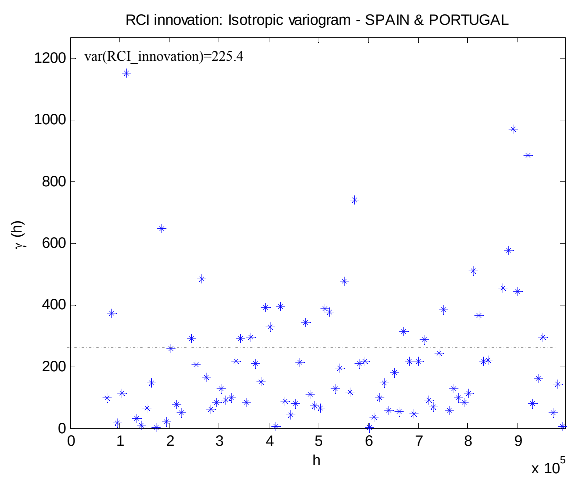
a) RCI total



b) RCI_basic



c) RCI_efficiency



d) RCI_innovation

Figure 19: Semi-variogram estimation for regions in Spain and Portugal – Euclidean distances

1.2.2 Distances along the road network

To make the analysis closer to real life, a second scenario is computed using as distances the length of the road (or ferry) which connects the regions along the actual network. Data are from the TRANSTOOLS road network tool (http://.energy.jrc.ec.europa.eu/transtools). The spatial correlation analysis with road distances should be more representative of the real connections across regions. As before, the analysis is carried out for countries in group A and B.

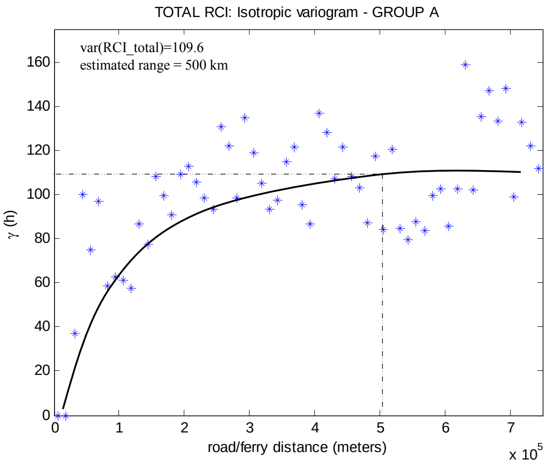
Table 3 shows the settings of the analysis in this case and main results.

Table 3: Basic settings for the semi-variogram analysis and estimated ranges with road/ferry distances between regions

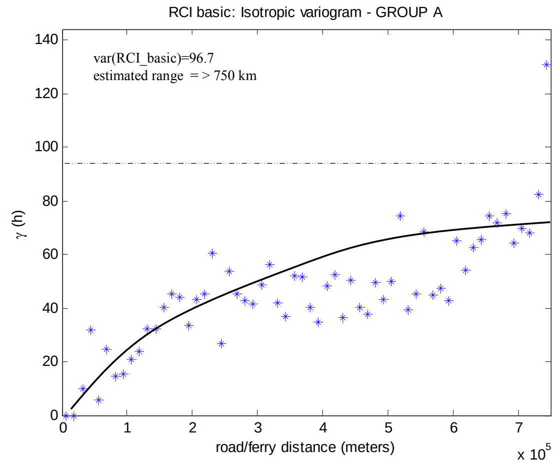
	MAXIMUM ROAD DISTANCE BETWEEN CENTROIDS§	MAXIMUM ROAD DISTANCE FOR VARIOGRAM ESTIMATION	variable of interest (min-max normalized values)	variance (asymptothic value for the semi-variogram)	estimated range (km)
GROUP A	3721 km	750 km	RCI_total	109.6	500
			RCI_basic	96.7	> 750
			RCI_efficiency	91.5	350
			RCI_innovation	190.5	350
GROUP B	4040 km	750 km	RCI_total	251.7	na
			RCI_basic	223.3	>750
			RCI_efficiency	178.7	na
			RCI_innovation	228.9	na

§ Maximum distance between region centroids on road or ferry networks

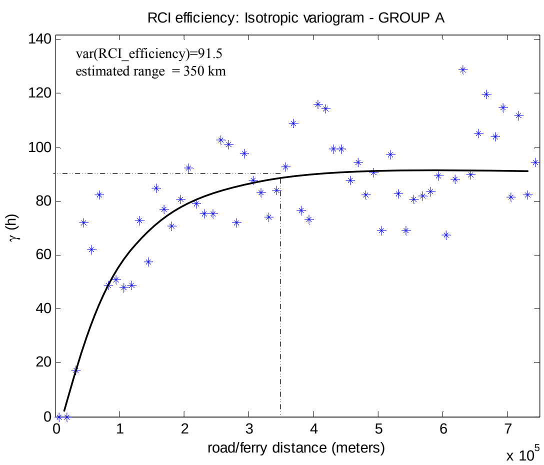
Estimated ranges are very similar to the ones estimated using Euclidean distances (Table 2). Figure 20 and Figure 21 show the sample points and the approximate estimated variograms (solid black line) for the two groups of countries. Results for efficiency and innovation sub-indices for group A indicate a rather short 'spill-over' effect with ranges, along the road networks, of about 350 km. Also in this case the analysis shows a clear covariance structure for the countries in group A, while countries in group B are not showing any spatial autocorrelation.



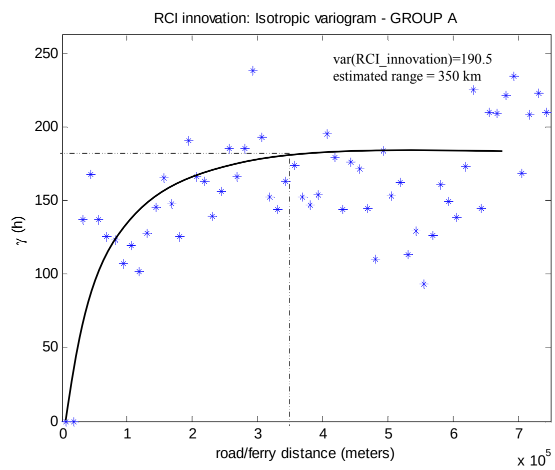
a) RCI total



b) RCI_basic

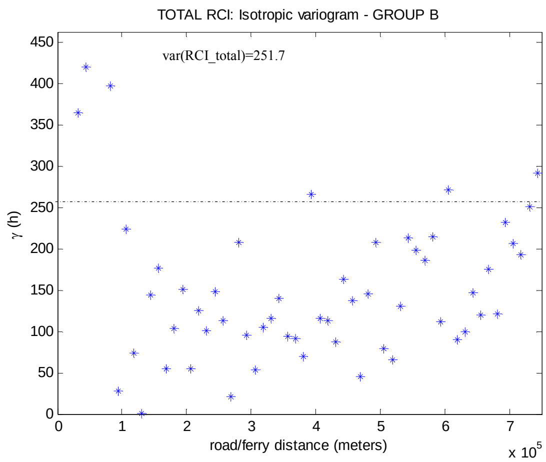


c) RCI_efficiency

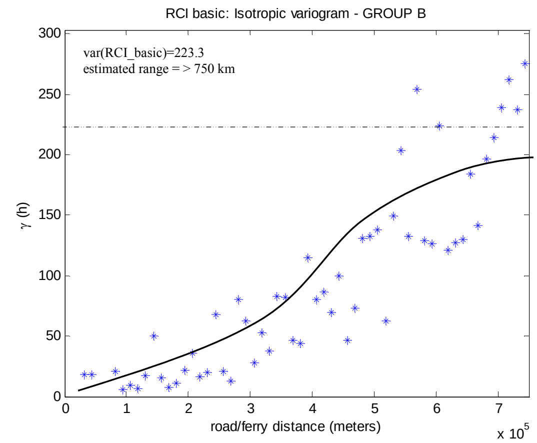


d) RCI_innovation

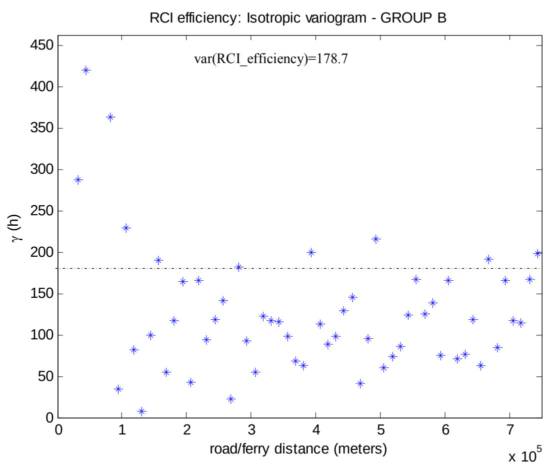
Figure 20: Semi-variogram estimation for regions in group A – road distances



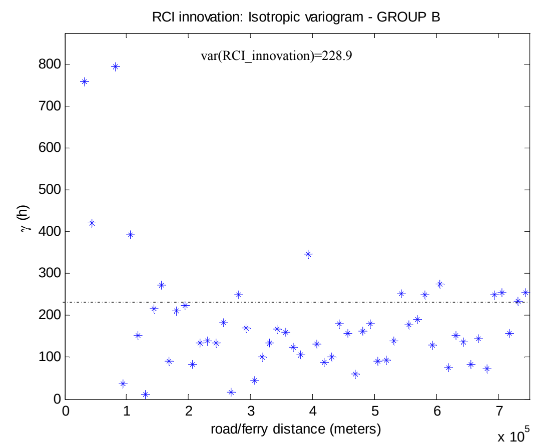
a) RCI total



b) RCI_basic



c) RCI_efficiency



d) RCI_innovation

Figure 21: Semi-variogram estimation for regions in group B – road distances

1.2.3 Travel-time distances

The previous analyses suggest a deeper look into the first group of countries, which are the only ones showing a clear covariance structure. As a third scenario we estimated the variograms for countries in group A using travel-time distances, which are a better proxy of the actual connectivity between regions. Travel-time distances are estimated between population-weighted centroids of NUTS2 regions, using the TRANSTOOLS road network as in the previous case (http://energy.jrc.ec.europa.eu/transtools/).

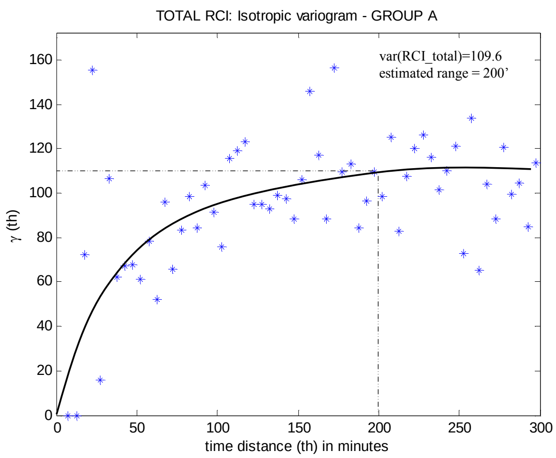
Table 4: Basic settings for the semi-variogram analysis and estimated ranges with travel-time distances between centroids

	MAXIMUM TIME DISTANCE BETWEEN CENTROIDS	MAXIMUM TIME DISTANCE FOR VARIOGRAM ESTIMATION	variable of interest (min-max normalized values)	variance (asymptothic value for the semi-variogram)	estimated range (minutes)
GROUP A	40 hours	5 hours (300')	RCI_total	109.6	200'
			RCI_basic	96.7	> 300'
			RCI_efficiency	91.5	150'
			RCI_innovation	190.5	150'

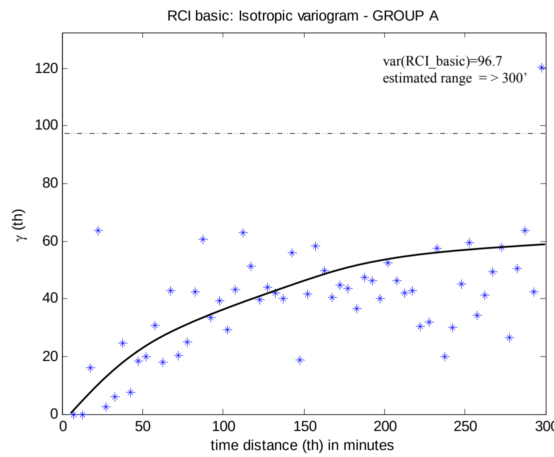
Basic settings for the variogram estimation with travel-time distances are shown in Table 4. The maximum travel-time distance is set to 300 minutes (5 hours). Asymptotic values of the variograms are the same as the previous case, since they are the sample variances of the four indices over the two sub-areas.

Figure 22 shows the estimated variograms with time-ranges which go from a minimum of 150 minutes, for the Innovation sub-index (Figure 22-d), to a maximum of 200 minutes, for the RCI total index(Figure 22-a). The time-range for the RCI-basic (Figure 22-b) is higher than 300 minutes, the maximum travel time distance set for the computations, and it is not of no relevance for the analysis as this index is mostly at the country level. As for the previous case, efficiency and innovation sub-indices show a clear correlation structure with a maximum correlation (time) distance of about 2.5 hours.

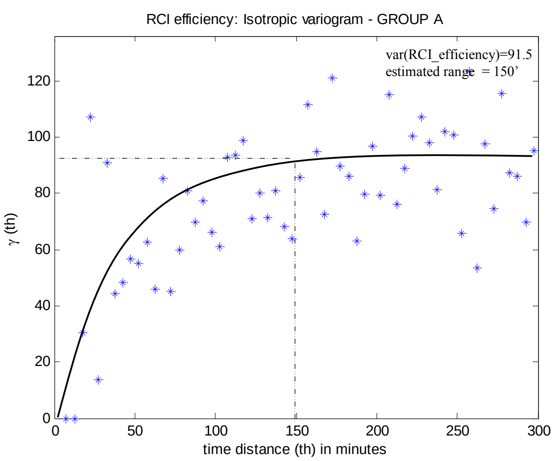
Anomalous points can be detected in the upper-left corner of Figure 20 a-d and Figure 22 a-c-d. A deeper look at the data highlighted that this particular estimation of $[h,\gamma(h)]$ is associated to pairs of regions which scored quite different values of RCI indices even if they are very close to each other - about 40 km along the road network or 20-25 minutes of travel time. These pairs are for example AT12-AT13 (Niederösterreich-Wien); BE31-BE35 (Brabant Wallon-Namur) and UKD2-UKD5 (Cheshire-Merseyside).



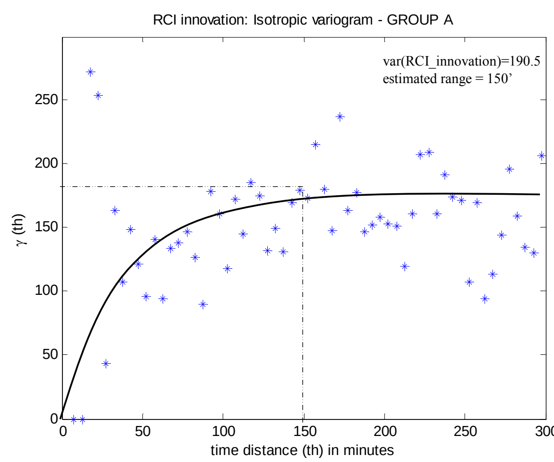
a) RCI total



b) RCI_basic



c) RCI_efficiency



d) RCI_innovation

Figure 22: Semi-variogram estimation with travel-time distances for regions in group A

In conclusion, outcomes of the variogram analysis indicate the existence of a clear structure of correlation for the sub-area A of high-high clusters. In this area the range of auto-correlation is between 300-500 km for Euclidean and road distance, while in terms of travel-time distance the estimated range is about 150 and 200 minutes. Variogram analysis cannot estimate a range for the sub-area of low-low clusters, sub-area B. This area seems to be characterized mostly by low performing regions with some rare and sparse picks of relatively higher performers (some capital regions).

2 Association with exogenous indicators

The Regional Competitiveness Index comprises many different factors, spanning from the quality of Institutions to the level of Innovation of the regional economy. Nevertheless two types of aspects are not directly included in the Index. They are: a. factors related to demographic change and b. factors related to economic growth.

The aim of this part is to explore possible significant associations between these aspects and RCI indices by means of a bivariate correlation analysis. Three separate analyses are carried out: one considering all the European regions and the other two separately considering the two sub-areas – A and B –detected by the spatial autocorrelation analysis discussed in Section 1.

Information related to the population is indirectly included in the Index as almost all the indicators are normalized to the number of inhabitants. The only exceptions are the indicators which describe the market size. Plus, all the indicators are standardized by using weighted mean and standard deviation, with weights being the share of population. Other demographic factors, especially related to population trends, have not been taken directly into account in the Index. Economic indicators are instead the backbone of the RCI, but they are all measured at one point in time (the latest available year) and not as rates or trends. With all the limitations imposed by the fact that RCI covers a period which includes the 2008 economic and financial crisis, some interesting relationships may arise between RCI scores and a proxy for regional economic growth.

To this twofold aim, a simple correlation analysis is carried out between RCI scores and the following indicators:

1. population change in the period 2001-2007;
2. natural population change in the period 2001-2007;
3. net migration in the period 2001-2007;
4. share of population which live in Large Urban Zones, LUZ;
5. GDP growth average 2000-2007;

The first three indicators describe main population trends in terms of growth rates due to natural changes or migration flows. Population change is based on the difference between the population on 1st January of year *t*-1 and of 1st of January of year *t*. Natural population change is computed as the number of births minus the number of deaths in a calendar year; while net-migration is the residual, i.e. the difference between the two population changes. The fourth indicator is a proxy of the spatial distribution of the population while the fifth indicator is included as a proxy of economic growth.

The analysis of correlation is carried out between these five exogenous indicators and RCI total score, RCI_eff and RCI_inn, sub-indices for efficiency and innovation.

Table 5: Correlation coefficients between RCI scores and exogenous indicators. Values in bold are statistically significant at the level α = 0.05.

all countries sample size N =268
critical value for N> 100 at level 0.05 = 0.195

	RCI_total	RCI_eff	RCI_inn
population change 01-07	0.11	0.05	0.17
natural population change 01-07	0.11	0.01	**0.25**
net migration 01-07	0.07	0.06	0.06
share of population in LUZ	**0.41**	**0.44**	**0.42**
GDP growth average 00-07	**-0.31**	-0.17	**-0.36**

group A sample size N =128
critical value for N> 100 at level 0.05 = 0.195

	RCI_total	RCI_eff	RCI_inn
population change 01-07	**0.24**	**0.29**	0.16
natural population change 01-07	**0.52**	**0.50**	**0.40**
net migration 01-07	-0.04	0.04	-0.07
share of population in LUZ	**0.21**	**0.22**	**0.29**
GDP growth average 00-07	**-0.21**	-0.15	**-0.30**

group B sample size N = 65
critical value for N = 60 at level 0.05 = 0.25

	RCI_total	RCI_eff	RCI_inn
population change 01-07	**0.33**	**0.23**	**0.39**
natural population change 01-07	**0.27**	0.13	**0.26**
net migration 01-07	**0.26**	**0.22**	**0.34**
share of population in LUZ	**0.76**	**0.79**	**0.73**
GDP growth average 00-07	**0.25**	**0.46**	**0.36**

Correlation coefficients are shown in Table 5 where values in bold indicate coefficients statistically significant at the level α = 0.05.

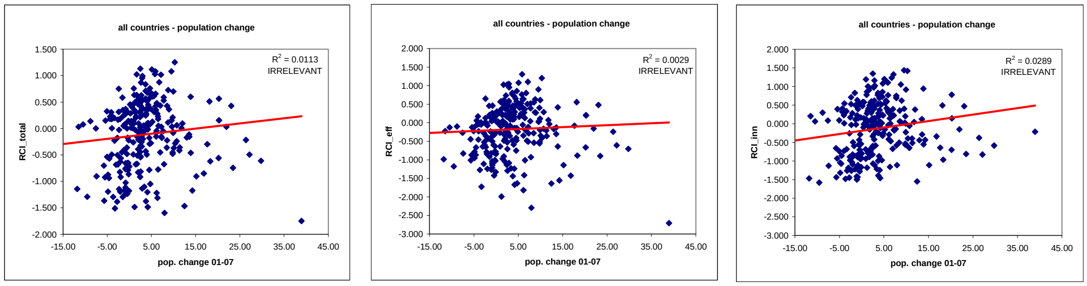
It is interesting to note that the number of significant relationships increases by separately considering the two sub-areas A and B. When considering all the European regions (Table 5 – top box) significant associations are found only for 6 cases out of 15, while 10 and 14 cases out of 15 are significant respectively for sub-area A and B. For each pair, Figure 23-Figure 27 show the scatter plot with the corresponding linear regression and its coefficient of determination R^2.

The share of population in LUZ is always positively associated to the three RCI indices (total, efficiency and innovation), with particularly high values for countries in sub-area B. In these regions living in high density areas (large cities) means having higher levels of competitiveness. On the contrary, GDP growth average shows an anomalous negative association with RCI_total and RCI_inn for the analysis including all regions and for the one in sub-area A. These results are not true for regions in sub-area B where the association between GDP growth average and RCI_indices is always significant and positive. This may be due to the fact that the two indicators cover different time periods - GDP growth average is computed for the period 2000-2007, while the RCI data span the period 2007-2009, including also the 2008 economic and financial crisis. In fact, for some dimensions of competitiveness RCI already reflects the negative effects of the crisis and it offers a much more complex picture of competitiveness including a number of indicators not strictly related to economic performance. We can have more reliable results for the association between GDP growth and the RCI once we can analyze the GDP growth for the same time period as covered by the data included in the RCI. Besides, it is worth noting that the relationship between competitiveness and growth is in general difficult to understand. A recent example is the Trade Performance Index, jointly developed by UNCTAD (United Nations Conference on Trade and Development) and WTO (World Trade Organization), which measures level of competitiveness and diversification of the export sectors of about 180 countries. The most competitive countries in terms of Trade Performance Index scores are those where GDP growth was the lowest in the last ten years, while low performing countries are those where the GDP growth has been the highest in the last decade (Fortis, 2010).

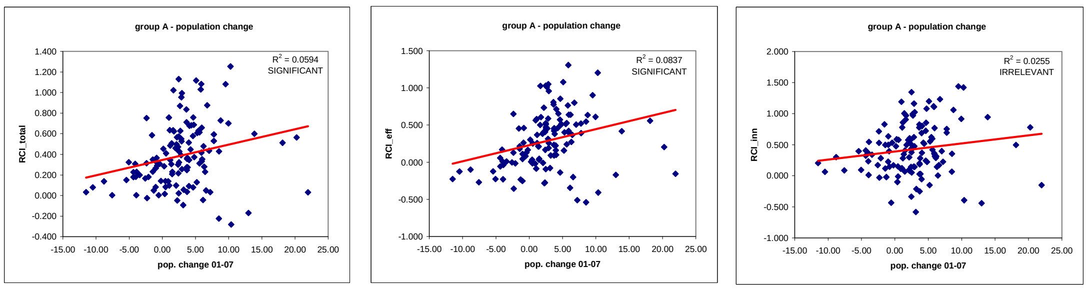
Focusing on countries in sub-area A, where spatial auto-correlation analysis highlighted clusters of highly performing regions, all the indicators but net migration and GDP growth average are positivity correlated with RCI indices (the only irrelevant correlation is between population change and RCI_inn). The “net migration” gives always irrelevant correlation while the anomalous behavior of GDP growth average for these regions has just been discussed. Other significant and direct associations in sub-area A are observed for population change and natural population change. This is in line with the assumption that population dynamics do matter for territorial competitiveness.

The analysis for regions in sub-area B shows a positive correlation for all the exogenous indicators and almost all the RCI indices (with the only exception of RCI_eff and natural population change). In this area, where the spatial analysis highlighted clusters of low performing regions, demographic trends are highly relevant for territorial competitiveness. It is worth noting that the analysis of correlation is not a causal-effect analysis. So, in this particular case, it is not possible to distinguish whether countries with stronger population dynamics in terms of population change, inflows and outflows stimulate competitiveness or competitive environments help demographic vivacity. The point is that there exists a positive strong association between the two factors which is stronger for regions in sub-area B, with clusters of low performers and some isolated picks of high performers, than for those in sub-area A, with more homogeneous clusters of high performers.

ALL COUNTRIES ANALYSIS



COUNTRIES IN GROUP A: AT, BE, DE, DK, FI, IE, LU, NL, SE, SI, UK



COUNTRIES IN GROUP B: BG, CZ, EE, GR, HU, LT, LV, PL, RO, SK

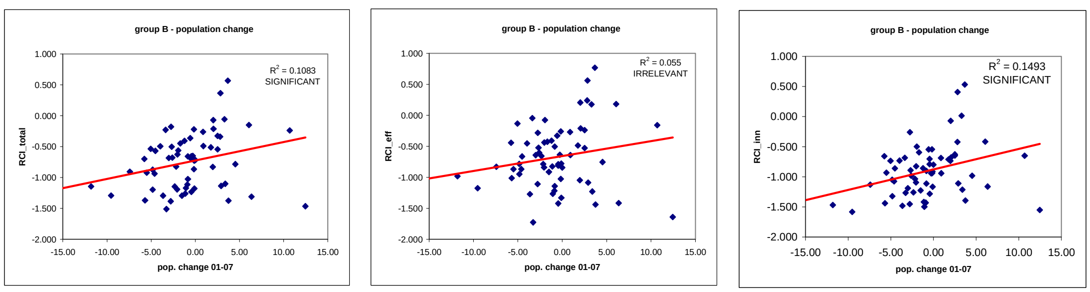
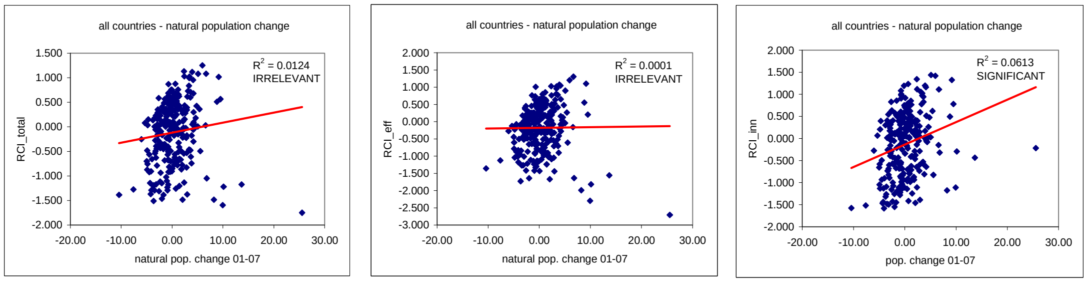
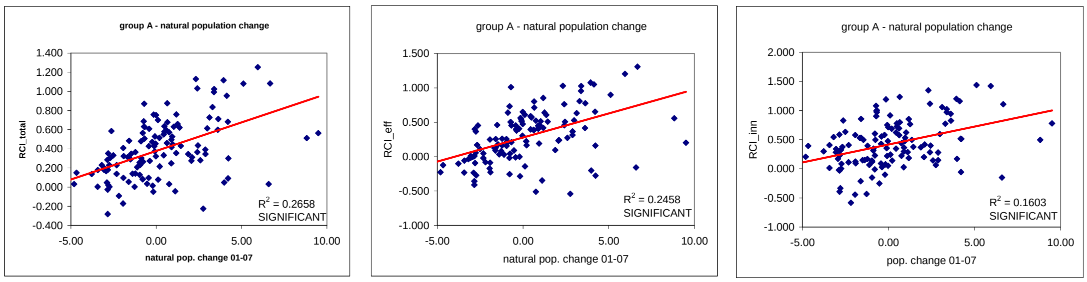


Figure 23: Association between RCI indices and population change 01-07

ALL COUNTRIES ANALYSIS



COUNTRIES IN GROUP A: AT, BE, DE, DK, FI, IE, LU, NL, SE, SI, UK



COUNTRIES IN GROUP B: BG, CZ, EE, GR, HU, LT, LV, PL, RO, SK

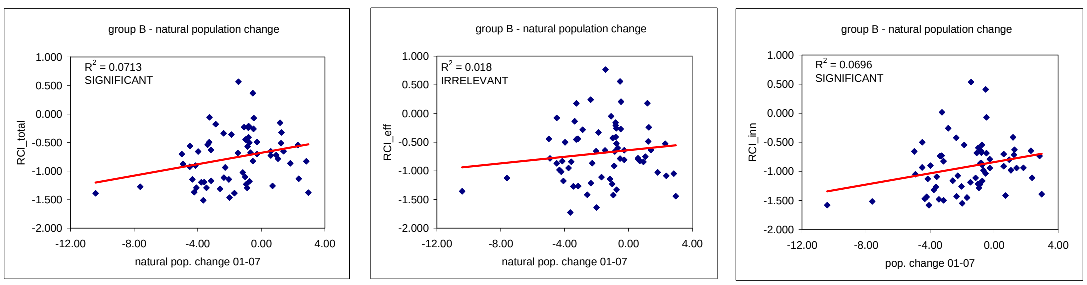
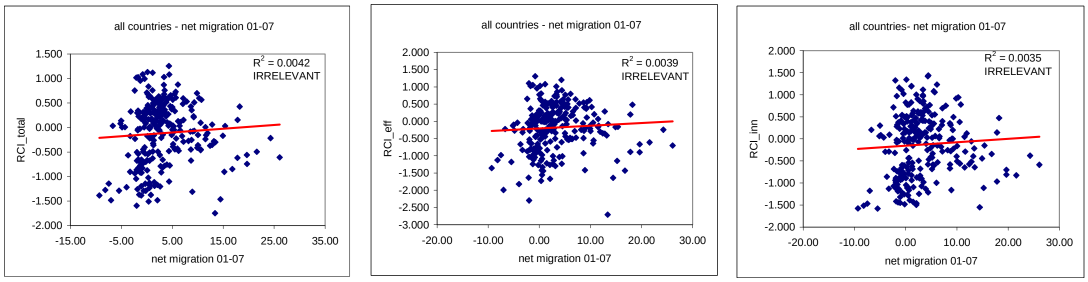
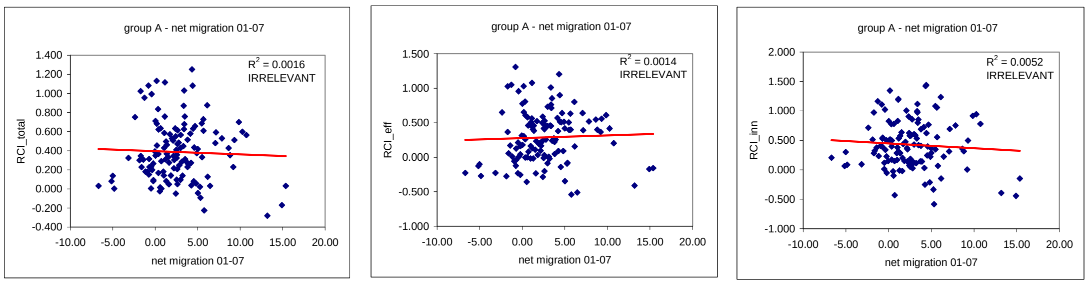


Figure 24: Association between RCI indices and natural population change 01-07

ALL COUNTRIES ANALYSIS



COUNTRIES IN GROUP A: AT, BE, DE, DK, FI, IE, LU, NL, SE, SI, UK



COUNTRIES IN GROUP B: BG, CZ, EE, GR, HU, LT, LV, PL, RO, SK

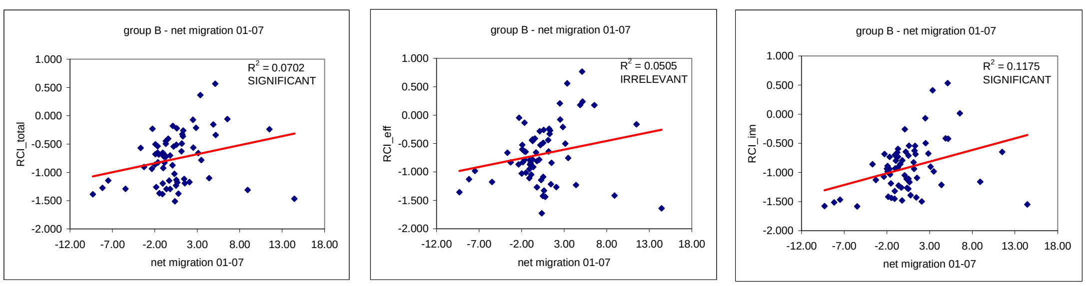


Figure 25: Association between RCI indices and net migration 01-07

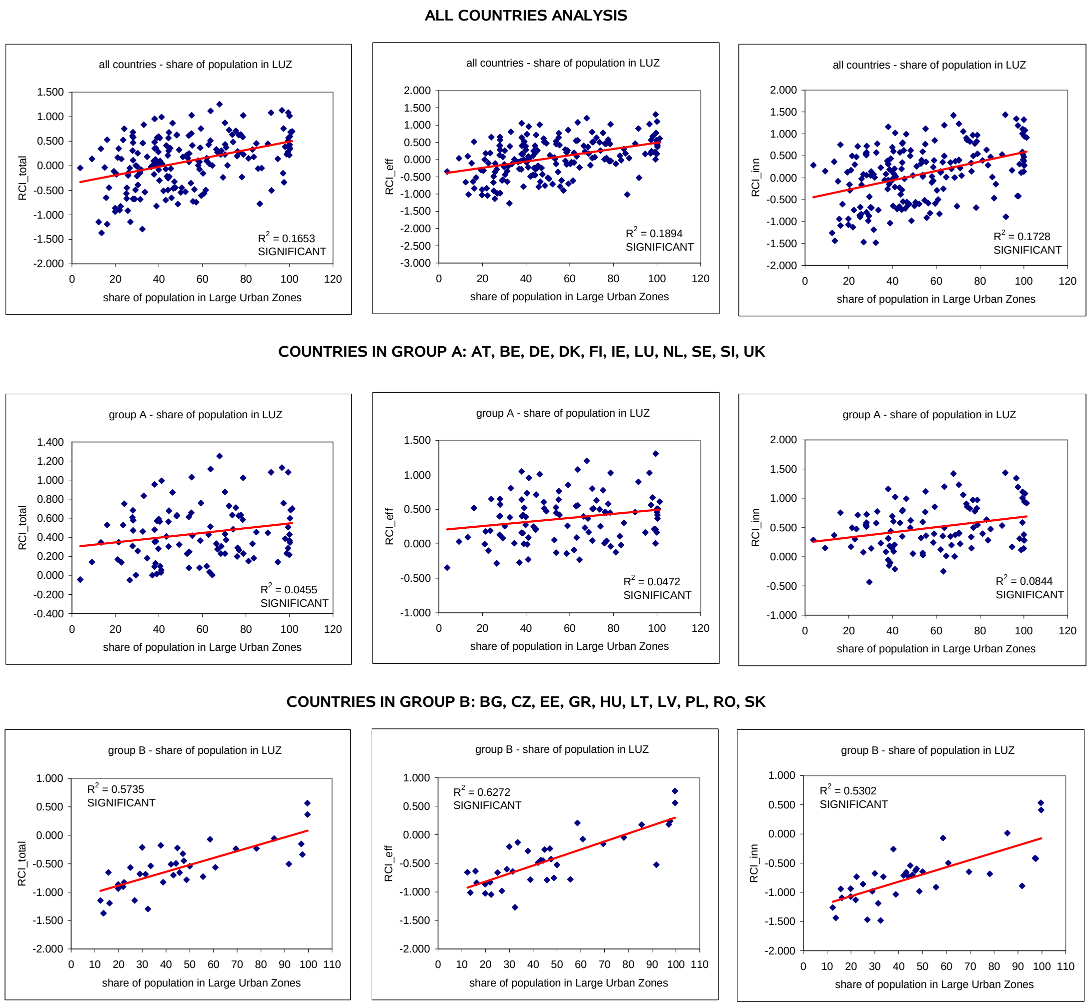
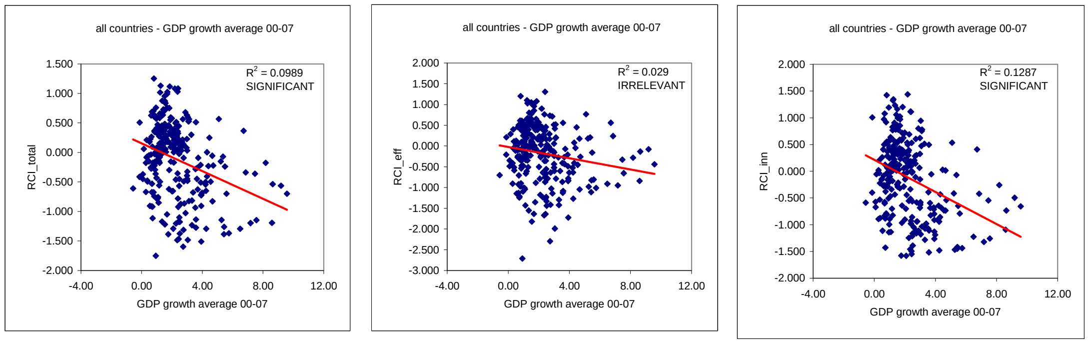
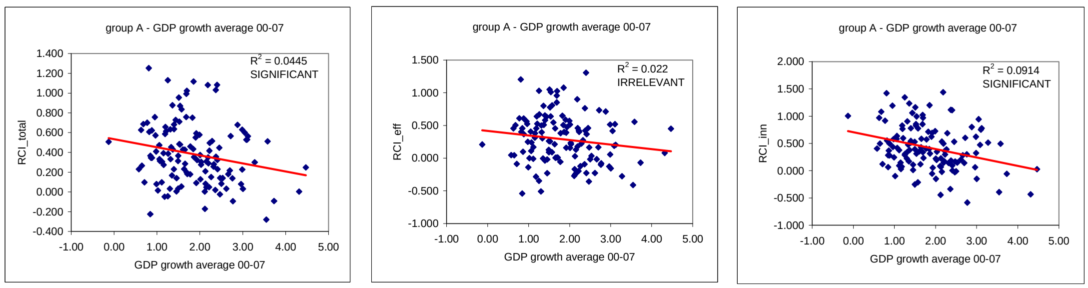


Figure 26: Association between RCI indices and share of population in Large Urban Zones

ALL COUNTRIES ANALYSIS



COUNTRIES IN GROUP A: AT, BE, DE, DK, FI, IE, LU, NL, SE, SI, UK



COUNTRIES IN GROUP B: BG, CZ, EE, GR, HU, LT, LV, PL, RO, SK

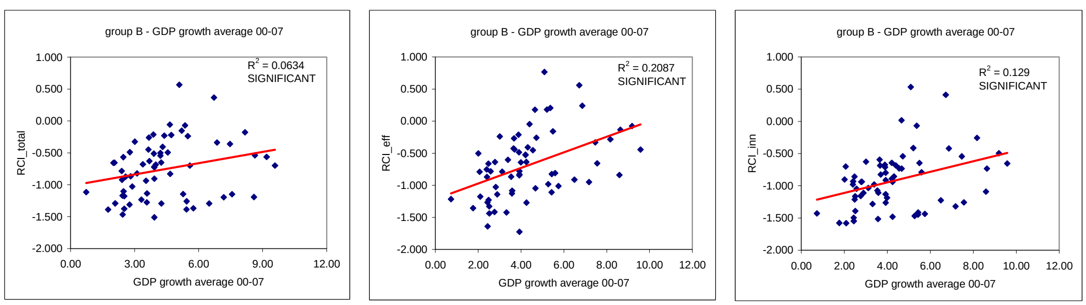


Figure 27: Association between RCI indices and GDP growth average 00-07

Conclusions

Outcomes from the ESDA analysis of the RCI suggest the existence of regional spillovers among EU NUTS 2 regions proxied by global spatial autocorrelation. Specifically, there is evidence of significant spatial correlation and potential spillover effect among regions in Bulgaria, Romania and Greece, all with low RCI performance. On the other hand, regional spillovers from strong competitiveness performance are observed in the Southern part of the UK, parts of Belgium, the Netherlands, Germany, Denmark and a few Scandinavian regions. We can clearly observe that the regions with strong competitiveness which show evidence of spillover effects constitute the upper part of the blue banana or the European regions which traditionally score very high on a number of economic indicators such as productivity, innovative capacity, use of new technology.

Our variogram analysis is separately carried out for 'homogenous' sub-areas as detected by the ESDA analysis. Outcomes show the existence of a clear structure of spatial auto-correlations in the sub-area with clusters of high-high regions. In this sub-area the maximum range of action of the spillover effects ranges from 300-500 km when taking into account two different distances between regions (Euclidean and distance along the real road network) and between 150 and 200 minutes in terms of travel-time distance. Variogram analysis cannot estimate a range for the sub-area of low-low clusters and in some other areas comprising Italy, France, Spain and Portugal. These areas do not show clear structures of spatial auto-correlation save for clusters of low-low regions or very heterogeneous pictures.

Anselin et al. (2007) point out that spatial analysis can serve as a useful tool for countries to monitor social indicators but it is important to keep in mind that these exploratory techniques are only suggestive of possible hypotheses and relations. The main contribution of spatial autocorrelation analysis is to highlight potentially interesting features in the data, and to facilitate the discovery process (Anselin et al. 2007). It is, however, important to look at the spatial distribution and presence of clusters and outliers for relevant indicators at the EU level as interesting patterns can be revealed and more informed policy decisions could be undertaken, especially at the regional level. In the future, when the new edition of the RCI is prepared, temporal analysis can be added to the spatial one.

With regards to the analysis of analysis of possible relationships between exogenous indicators and the RCI score and sub-scores, we find that the number of significant results from the correlation analysis increase significantly when we distinguish among different sub-areas within the EU territory, which show less heterogeneity in RCI indices. The share of population living in LUZ is always positively associated to the three RCI indices (total, efficiency and innovation), with particularly high values for countries in sub-area B. In these regions living in high density areas (large cities) means having higher levels of competitiveness. Focusing on countries in sub-area A, where spatial auto-correlation analysis highlighted clusters of highly performing regions, all the indicators but net migration and GDP growth average are positively correlated with RCI indices The analysis for regions in sub-area B shows a positive correlation for all the exogenous indicators and almost all the RCI indices (with the only exception of the efficiency sub-index and natural population change). In general, our results show that population dynamics and demographic trends are highly relevant for territorial competitiveness. On the contrary the relationship between RCI indices and GDP growth is unclear. This is indeed in line with some recent discussions about the difficulty of interpretation of the relation between competitiveness/productivity and economic growth. A popular example is the ambiguous link between the last release of the Trade Performance Index by UNCTAD/WTO and GDP growth in the last decade, where most competitive countries are those which show the lowest GDP growth and least competitive have experienced the highest growth (Fortis, 2010).

References

Annoni, P. and K. Kozovska (2010). "EU Regional Competitiveness Index (RCI) 2010." JRC Scientific and Technical Report EUR 24346 EN. Luxembourg: Publications Office of the European Union.

Anselin, L., Sridharan, S. and S. Gholston. (2007) "Using exploratory spatial data analsyis to leverage social indicator databases: the discovery of interesting patterns." *Social Indicators Research*, 82: 287-309.

Anselin, L., Syabri, I. and Y.Kho (2006). "GeoDa: An Introduction to Spatial Data Analysis." *Geographical Analysis*, 27:93-115.

Anselin, L., Kim, Y.-W., and Syabri, I. (2004). "Web-based spatial analysis tools for the exploration of spatial data." *Journal of Geographical Systems*, 6.

Anselin, L., Syabri, I., and Smirnov, O. (2002). "Visualizing multivariate spatial correlation with dynamically linked windows." In Anselin, L. and Rey, S. (eds) *New Tools for Spatial Data Analysis: Proceedings of the Specialist Meeting*. Center for Spatially Integrated Social Science (CSISS), University of California, Santa Barbara.

Anselin, L. (1999) "Interactive techniques and exploratory spatial data analysis." In Longley, P., et al. (eds) *Geographical Information Systems: Principles, Techniques, Management and Applications*. New York: Wiley.

Anselin, L. (1996). "The Moran scatterplot as an ESDA tool to assess local instability in spatial association." In Fischer, M., Scholten, H., and Unwin, D. (eds) *Spatial Analytical Perspectives on GIS in Environmental and Socio-Economic Sciences*. London: Taylor and Francis.

Anselin, L. (1995). "Local indicators of spatial association — LISA." *Geographical Analysis*, 27:93–115.

Anselin, L. (1988). *Spatial econometrics: methods and models*. Dordrecht: Kluwer.

Baumont, C., Ertur, C. and J. Le Gallo (2003). "A spatial econometric analysis of geographical spillovers and growth for European regions, 1980-1995." In Fingleton, B. *European Regional Growth.* Heidelberg: Springer Verlag.

Cressie, N. (1984) "Towards resistant geostatistics." In Verly, G. et al (eds) *Geostatistics for Natural Resources Characterization*. Dordrecht: Reidel.

Dall'Erba, S. and J. Le Gallo (2008) "Regional convergence and the impact of European structural funds over 1989-1999: a spatial econometric analysis." *Papers in Regional Science*, 87(2): 219-244.

de Marsily G. (1986) "Geostatistic and Stochastic Approach in Hydrogeology." In *Quantitative Hydrogeology*, chapter 11, 286-329, Academic Press INC.

Ertur, C. and W. Koch (2006) "Regional disparities in the European Union and the enlargement process: an exploratory spatial data analysis, 1995–2000." *Annals of Regional Science* 40:723–765.

Fortis, M. (2010) "La competitività del sistema produttivo italiano: effetto statistico o realtà economica?" Introductory speech at the X^{th} National Conference of Statitistics - ISTAT, Rome December, 15-16 2010.

Getis, A. and J. Aldstadt (2002) "Constructing the spatial weights matrix using a local statistic." *Geographical Analysis*, 34 (2): 130-140.

Haining, R. (2003) *Spatial Data Analysis. Theory and Practice*. Cambridge: Cambridge University Press.

Islam, N. (2003) “What have we learnt from the convergence debate?” *Journal of Economic Surveys* 17, 309–362.

Le Gallo, J. and C. Ertur (2003) “Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe, 1980–1995.” *Papers in Regional Science* 82: 175–201.

Magrini, S. (2004) “Regional (di)convergence.” In: Henderson, V., Thisse, J.F. (eds.), *Handbook of Regional and Urban Economics*, vol. 4. Amsterdam: Elsevier.

Moran P.A.P. (1950) “Notes on continuous stochastic phenomena.” *Biometrika*, 37:17-23

Perry G.L., Miller, B.P., Enright, N.J. (2006) “A comparison of methods for the statistical analysis of spatial point patterns in plant ecology.” *Plant Ecology,* 187: 59-82.

Ramajo, J.. Marquez, M., Hewings, G. and M. Salinas (2008) “Spatial heterogeneity and interregional spillovers in the European Union: Do cohesion policies encourage convergence across regions?” *European Economic Review*, 52: 551-567.

Thompson S. K. (1992) *Sampling*. Wiley-Interscience Publication.

Appendix A

LISA statistics for RCI_total and associated signicance levels

code	LISA	
BE10	HH	***
BE21	HH	***
BE22	HH	***
BE23	HH	***
BE24	HH	***
BE25	HH	
BE31	HH	**
BE32	HH	***
BE33	HH	
BE34	LH	
BE35	HH	**
BG31	LL	***
BG32	LL	***
BG33	LL	***
BG34	LL	***
BG41	LL	***
BG42	LL	***
CZ01	HL	
CZ02	LH	
CZ03	LH	
CZ04	LH	
CZ05	LL	
CZ06	LH	
CZ07	LL	
CZ08	LL	
DK01	HH	**
DK02	HH	***
DK03	HH	***
DK04	HH	***
DK05	HH	***
DE11	HH	**
DE12	HH	**
DE13	HH	**
DE14	HH	**
DE21	HH	
DE22	HH	
DE23	HH	
DE24	HH	
DE25	HH	
DE26	HH	**
DE27	HH	**
DE30	HH	
DE41	HL	
DE42	HH	
DE50	HH	
DE60	HH	
DE71	HH	**
DE72	HH	
DE73	HH	
DE80	HH	
DE91	HH	
DE92	HH	
DE93	HH	
DE94	HH	

code	LISA	
DEA1	HH	***
DEA2	HH	**
DEA3	HH	***
DEA4	HH	
DEA5	HH	**
DEB1	HH	**
DEB2	HH	
DEB3	HH	**
DEC0	HH	
DED1	HH	
DED2	HH	
DED3	HH	
DEE0	HH	
DEF0	HH	
DEG0	HH	
EE00	LH	
IE01	HH	
IE02	HH	
GR11	LL	***
GR12	LL	***
GR13	LL	***
GR14	LL	***
GR21	LL	***
GR22	LL	***
GR23	LL	***
GR24	LL	***
GR25	LL	***
GR30	LL	***
GR41	LL	***
GR42	LL	***
GR43	LL	***
ES11	LL	
ES12	LL	
ES13	LL	
ES21	HL	
ES22	LL	
ES23	LL	
ES24	LH	
ES30	HL	**
ES41	LL	
ES42	LL	
ES43	LL	
ES51	HL	
ES52	LL	
ES53	LL	
ES61	LL	***
ES62	LL	
ES63	LL	***
ES64	LL	**
ES70	LL	***
FR10	HH	
FR21	LH	
FR22	HH	
FR23	HH	

code	LISA	
FR24	HH	
FR25	LH	
FR26	LH	
FR30	HH	
FR41	HH	
FR42	HH	
FR43	LH	
FR51	HH	
FR52	HH	
FR53	LL	
FR61	LH	***
FR62	HL	
FR63	LH	
FR71	HL	
FR72	LH	
FR81	LH	
FR82	HL	
FR83	LL	
ITC1	LH	***
ITC2	LH	
ITC3	LL	
ITC4	HL	
ITD1	LH	
ITD2	LH	
ITD3	HL	
ITD4	LL	
ITD5	HL	
ITE1	LL	
ITE2	LL	
ITE3	LL	
ITE4	HL	
ITF1	LL	
ITF2	LL	
ITF3	LL	**
ITF4	LL	**
ITF5	LL	**
ITF6	LL	**
ITG1	LL	**
ITG2	LL	
CY00	LL	***
LV00	LL	
LT00	LL	**
LU00	HH	
HU10	HL	**
HU21	LL	
HU22	LH	
HU23	LL	
HU31	LL	**
HU32	LL	***
HU33	LL	***
MT00	LL	**
NL11	HH	**
NL12	HH	***
NL13	HH	**

code	LISA	
NL21	HH	***
NL22	HH	***
NL23	HH	***
NL31	HH	***
NL32	HH	***
NL33	HH	***
NL34	HH	***
NL41	HH	***
NL42	HH	***
AT11	HH	
AT12	HH	
AT13	HH	
AT21	HH	
AT22	HL	
AT31	HH	
AT32	HH	
AT33	HH	
AT34	HH	
PL11	LL	
PL12	HL	***
PL21	LL	**
PL22	LL	
PL31	LL	
PL32	LL	
PL33	LL	
PL34	LL	
PL41	LL	**
PL42	LH	
PL43	LH	
PL51	LL	
PL52	LL	
PL61	LL	
PL62	LL	
PL63	LH	**
PT11	LL	
PT15	LL	**
PT16	LL	
PT17	HL	**
PT18	LL	
PT20	LL	**
PT30	LL	**
RO11	LL	***
RO12	LL	***
RO21	LL	***
RO22	LL	***
RO31	LL	***
RO32	LL	***
RO41	LL	***
RO42	LL	***
SI01	HH	
SI02	HH	
SK01	HH	
SK02	LL	
SK03	LL	

code	LISA	
SK04	LL	**
FI13	HH	
FI18	HH	**
FI19	HH	
FI1A	HH	**
FI20	HH	***
SE11	HH	
SE12	HH	**
SE21	HH	***
SE22	HH	**
SE23	HH	***
SE31	HH	**
SE32	HH	
SE33	HH	**
UKC1	HH	
UKC2	HH	
UKD1	LH	***
UKD2	HH	
UKD3	HH	
UKD4	HH	
UKD5	HH	**
UKE1	HH	
UKE2	HH	
UKE3	HH	
UKE4	HH	
UKF1	HH	
UKF2	HH	
UKF3	LH	
UKG1	HH	**
UKG2	HH	**
UKG3	HH	**
UKH1	HH	**
UKH2	HH	***
UKH3	HH	***
UKI1	HH	***
UKI2	HH	***
UKJ1	HH	***
UKJ2	HH	***
UKJ3	HH	***
UKJ4	HH	***
UKK1	HH	***
UKK2	HH	***
UKK3	LH	
UKK4	HH	
UKL1	HH	**
UKL2	HH	**
UKM2	HH	
UKM3	HH	
UKM5	HH	
UKM6	LH	***
UKN0	HH	

*** p<0.01, ** p<0.05

European Commission

EUR 24703 EN – Joint Research Centre – Institute for the Protection and Security of the Citizen
Title: RCI 2010: Some in-depth analysis
Author(s): Paola Annoni and Kornelia Kozovska
Luxembourg: Publications Office of the European Union
2011 – 55 pp. – 21.0 x 29.7 cm
EUR – Scientific and Technical Research series – ISSN 1018-5593
ISBN 978-92-79-19078-0
doi:10.2788/28476

Abstract
This document is the final delivery of the two-year joint project DG Joint Research Centre and DG Regional Policy on the measurement of the level of regional competitiveness, launched in November 2008. Within this project, the European Commission has recently published the first edition of the Regional Competitiveness Index (RCI). The index provides a tool to improve the understanding of competitiveness at the regional level by showing the strengths and weaknesses of each of the European regions at the NUTS2 level in a number of dimensions related to competitiveness. The analysis offered by the first edition of the RCI is a snapshot of regional competitiveness as it is in 2010 and is based upon data mostly spanning between 2007 and 2009. The present document takes a step further and offers a two-fold analysis based on the RCI indices: an exploratory spatial data analysis and an analysis of possible relationships between exogenous indicators and the RCI index and sub-indices.

The exploratory spatial data analysis shows the existence of spatial dependence among EU regions, with different patterns for different areas within the EU. This can be taken as an indication for the existence of spatial externalities among regions and, when observed for high performing regions, as evidence, or better, as necessary condition for spillover effects. LISA analysis allowed us to distinguish between two sub-areas in the EU: group A which comprises regions with high RCI performance surrounded by regions with similar strong competitive performance and group B, comprising low-performing regions surrounded by low RCI performers. The analysis has been extended to better explore the structure of spatial autocorrelation within the two main sub-areas – A and B - of low-low and high-high clusters as detected by LISA. The analysis of sub-area B is meant to further investigate the possible presence of ‘negative’ spillover effects where low performing regions negatively affect their neighbours.

Spatial autocorrelation structure is investigated by using variogram analysis, a tool typical of Kriging for describing spatial dependences. Variogram analysis provides as additional information the ‘range of action’ of spatial dependence, which is the maximum distance beyond which the correlation can be considered null. Variogram analysis is carried out using three different distances between region centroids: Euclidean distance, distance along the road (ferry) network and the travel time distance. Results indicate the existence of a clear structure of correlation for the sub-area A of high-high clusters. On the contrary, sub-area B seems to be characterized mostly by low performing regions with some rare and sparse picks of relatively higher performers (some capital regions).

With regards to the analysis of possible relationships between exogenous indicators and RCI index and sub-indices, we have looked at bivariate correlations with five exogenous indicators (population change in the period 2001-2007; natural population change in the period 2001-2007; net migration in the period 2001-2007; share of population which live in Large Urban Zones, LUZ; GDP growth average 2000-2007) for all EU NUTS 2 regions as well as for two sub-areas as identified by the ESDA analysis. We find that the number of significant results from the correlation analysis increases when we distinguish between the sub-areas. Results show that population dynamics and demographic trends are highly relevant for territorial competitiveness while the relationship with GDP growth remains ambiguous.

How to obtain EU publications

Our priced publications are available from EU Bookshop (http://bookshop.europa.eu), where you can place an order with the sales agent of your choice.

The Publications Office has a worldwide network of sales agents. You can obtain their contact details by sending a fax to (352) 29 29-42758.

The mission of the JRC is to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of EU policies. As a service of the European Commission, the JRC functions as a reference centre of science and technology for the Union. Close to the policy-making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.

LB-NA-24703-EN-C

EUROPEAN COMMISSION

ISBN 978-92-79-19078-0

Publications Office