European Research Studies Journal
Volume XXIV, Special Issue 2, 2021

pp. 503-512

Typology of Smart Specializations Across European Regions

Submitted 18/03/21, 1st revision 17/04/21, 2nd revision 23/05/21, accepted 30/06/21

Korneliusz Pylak[1], Piotr Oleszczuk[2], Przemysław Kowalik[3]

Abstract:

Purpose: *The aim of the research was to examine and develop a typology of smart specializations in European regions. Such a typology is essential to compare regions in terms of the adopted innovation strategies and to assess their performance after the implementation of the Europe 2020 strategy.*
Design/Methodology/Approach: *Since smart specializations refer to 82 industries, we applied principal component analysis (PCA) to reduce the number of industries on the one hand and to find a typology of a limited number of unique specializations on the other. Horn's parallel analysis indicated a maximum number of fifteen components (potential specializations in the typology).*
Findings: *After analyzing fifteen components defined in PCA, twelve pointed to meaningful and explainable specializations that can be grouped into five domains of the typology, tourism, ICT, health, transportation, and environment.*
Practical Implications: *Since countries and regions define their specializations very differently (narrowly or broadly), the typology developed in the study enables to articulate diverse specializations using one common language and to compare the performance of regions that have chosen one or more identical specializations.*
Originality/Value: *The existing literature lacks a common typology of smart specializations, which may be essential in the upcoming evaluation of the Europe 2020 strategy and the performance of regions after the 2014–2020 period.*

Keywords: *Smart specializations, cohesion policy, regions, Europe, evaluation, principal component analysis.*

JEL codes: *R58, O52.*

Paper Type: *Research paper.*

[1]*Department of Quantitative Methods in Management, Faculty of Management, Lublin University of Technology, Lublin, Poland, k.pylak@pollub.pl*
[2]*Department of Quantitative Methods in Management, Faculty of Management, Lublin University of Technology, Lublin, Poland, p.oleszczuk@pollub.pl*
[3]*Department of Quantitative Methods in Management, Faculty of Management, Lublin University of Technology, Lublin, Poland, p.kowalik@pollub.pl*

1. Introduction

The concept of 'smart specialization' (SS) was defined by the research team of Foray, David, and Hall (2009) in the framework of the European Research Area (ERA) and encompasses one or more scientific and technological areas in which a region can specialize, leveraging its greatest strategic potential (Foray, 2015; Morgan, 2015). This identification of an unspecified number of areas, arising from the entrepreneurial discovery process (Barca, 2009; Foray, 2014), yields a very different scope of specialization among European regions. Since the outcome of the process, the specialization ‘domain,’ also referred to as priority or priority areas (Gianelle *et al.,* 2020b), is not precisely specified (Mäenpää and Teräs, 2018), regions define these domains both narrowly and more broadly based on iterative entrepreneurial processes (Vallance, 2017) and their routines (Kroll, 2015). Therefore, the specializations may range from software development (Olomouc region in the Czech Republic) to information and communication technologies (North Rhine-Westphalia in Germany), from agricultural sciences (Bosnia and Herzegovina) to bioeconomy (Lublin region in Poland) and so forth. This causes problems in comparing and identifying common technological links between regions (Iacobucci, 2014).

In just over a decade of implementing the SS concept with a budget of over €80 billion for 2014–2020 (Uyarra, 2018), it has become arguably the largest innovation policy experiment in the world (Foray, 2015; Radosevic and Ciampi Stancova, 2018). Assessing the effectiveness of SS is also a vague and challenging task (Gianelle *et al.*, 2020a). Since the introduction of SS there has been several attempts to measure the regional performance and the effectiveness of the Cohesion Policy (Di Cataldo and Monastiriotis, 2020; Gianelle *et al.,* 2020a, 2020b; Iacobucci and Guzzini, 2016; McCann and Ortega-Argilés, 2016), but the results are not fully conclusive (Di Cataldo and Monastiriotis, 2020). Most likely, it will not be until several years after 2020 that the evaluation of the implementation of the Cohesion Policy and the SS concept can yield sound results, once the long-term effects emerge. Till then, the methodology of the evaluation needs to be developed and tested (Hassink and Gong, 2019).

This paper addresses this demand by providing a first attempt to develop a typology of SS that may be used in more rigorous measurements of smart specializations (Hassink and Gong, 2019). This typology allows each region to be assigned the same types of SS, regardless of how broadly a region defined each of its specializations. Once the typology is implemented, any analysis of SS performance will be feasible because the same measures will be used. In this way, it is possible to answer, for example, the question of how the specialization of *‘experience and culture-based tourism’* has resulted in economic growth in European regions.

To develop the typology, we employed principal component analysis to reduce the 82 industries assigned to the 1,346 specializations of European regions. Horn's parallel analysis indicated a maximum number of fifteen components (potential specializations in the typology). After analyzing fifteen components defined in PCA, twelve pointed to meaningful and explainable specializations that can be grouped into five domains of the typology: tourism, ICT, health, transportation, and environment. In each domain we obtained quite distinctive SS, and thus, for example, in the tourism domain we can indicate unsophisticated and traditional tourism, experience and culture-based tourism, creative industries, culture and tourism, and pure tourism and MICE (meetings, incentives, conferences and exhibitions).

The paper is structured as follows. After an introduction containing a literature review and the resulting purpose of the paper, we present the data and methods in the next section. In the third section, we present empirical results and a discussion of possible specializations in typology, while we conclude the last section with a summary of the typology domains.

2. Data and Methods

To analyze the typology of smart specializations, we use the regional innovation strategies (S3) of European countries and regions submitted to the Smart Specialization Platform of the Joint Research Centre (JRC) of the European Commission in Seville (https://s3platform.jrc.ec.europa.eu). The database contains 1,346 specializations from 246 countries and regions. Each specialization included in the S3 documents is typified by economic, scientific, and political domains. We use only economic domains in the study, leaving room to develop a typology of the other domains by future research. In the economic domain, each specialization is depicted with 82 industries by NACE rev. 2 codes, taking *1* if the specialization can be covered by the industry and *0* otherwise.

Since the number of industries (variables) in the current typology is large, we need to reduce the number of observed variables without significant loss of information. To this end, principal component analysis (PCA), widely used in economic research, can be applied. Although the use of PCA is sometimes considered controversial for binary data, it is mathematically sound and is equivalent to principal coordinate analysis with the Euclidean metric as a similarity measure (Jolliffe, 1986). At the same time, PCA allows us to detect structure in the relationships between variables. These new variables, the principal components (PC), are obtained as an orthogonal transformation of the input variables. The basic idea here is to link the correlated variables. In this way, a new set of uncorrelated (orthogonal) variables is obtained and, as a result, the objects under study can be classified in terms of PCs. The total variance is the sum of the variance of the PCs, the subsequent components are chosen to maximize the variance not explained by the previous one: the initial ones explain most of the variability. More details are given below.

Let the data set under analysis consist of r observations of n variables $X_1, X_2, \ldots, X_n$ and let $PC1, PC2, \ldots, PCk$, where $k \leq n$, be the PCs. In practice, two methods are used to determine the PCs: diagonalization of the covariance matrix C (eigenvalues, eigenvectors) and singular value decomposition (SVD). The components $PC1, PC2, \ldots, PCk$, are linear combinations of the observed variables, i.e.

$$PCj = \sum_{i=1}^{n} a_{i,j} \cdot X_i,\ j = 1, \ldots, k.$$

In the first method, the PCj coefficient vector $v_j = [a_{1,j}, \ldots, a_{n,j}]$ is the eigenvector of the covariance matrix C of the variables $X_1, X_2, \ldots, X_n$, corresponding to the eigenvalue λ_j, the j-th largest eigenvalue of C. The eigenvalues $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_k$ of the matrix C are the variances of the components $PC1, PC2, \ldots, PCk$. The fraction of variance explained by the PCj is equal to $\frac{\lambda_j}{\lambda_1 + \lambda_2 + \ldots + \lambda_k}$. In the second method, the component variances $\lambda_1, \ldots, \lambda_k$ are expressed by singular values. In our study, we use SVD-based PCA R package (Blighe and Lun, 2020).

The first step in PCA is to determine the number of components that will be used in the subsequent analysis. The goal is to keep the number of components as small as possible without losing too much information. Our decision is based on several methods: Horn's parallel analysis (Horn, 1965), elbow point finding (Thorndike, 1953), Marčenko-Pastur limit method (Marčenko and Pastur, 1967), and the Gavish-Donoho method (Gavish and Donoho, 2014). Horn's parallel analysis involves several iterations of mixing observations in rows to perform PCA on a permuted matrix. The Elbow point is the point at which the decreasing portion of the variance explained by successive PCs makes no significant contribution. The Marčenko-Pastur limit method uses the maximum variance that can be explained by fully random PCs to determine the number of PCs. The Gavish-Donoho method is based on minimizing the reconstruction error from PCs by singular values.

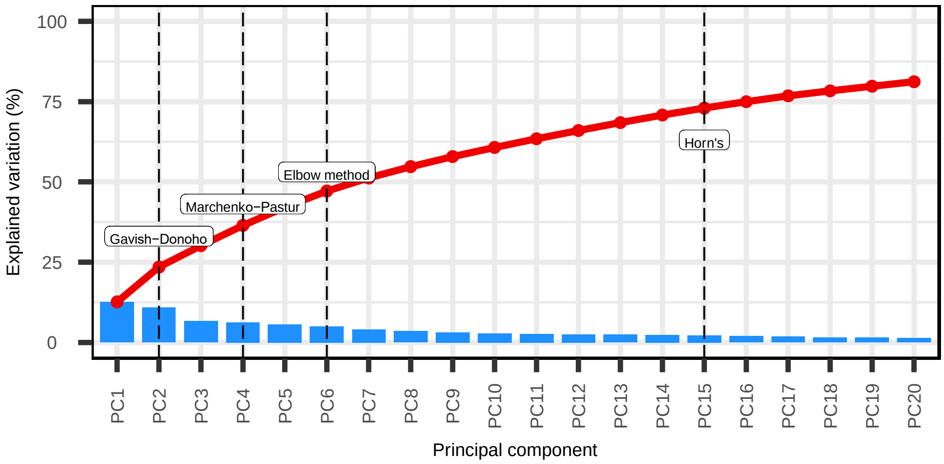
The second step involves analyzing the structure and relationships between the variables and the PCs. Since PCs are linear combinations of variables, the combination coefficients express the contribution or "loading" of individual variables to each principal component. The latter are called loadings and express the correlation coefficients of the observed variables and PCs. Based on the obtained loadings, we classify industries into components and prepare a typology of smart specializations in the following section.

3. Results and Discussion

The paper seeks to find unique smart specializations across European regions. The diversity in defining smart specializations causes difficulties in their analysis, as they cover different ranges of industries. By applying Principal Component Analysis to 1,436 smart specializations, it was possible to analyze the basic structure of specializations and transform the large number of industries (82) covered by each specialization into a lesser number of uncorrelated specialization typology.

Figure 1 presents the scree plot (Cattell, 1966) with the distribution of explained variance for the first twenty principal components. The vertical lines denote the different techniques for selecting the optimal number of principal components to retain, as described in the previous section. As indicated, the maximum number of components is 15 (performed by Horn's parallel analysis) and hence our discussion of the results will initially include this number of components.

Figure 1. *The distribution of explained variance across twenty first principal components of industrial domains in smart specializations*

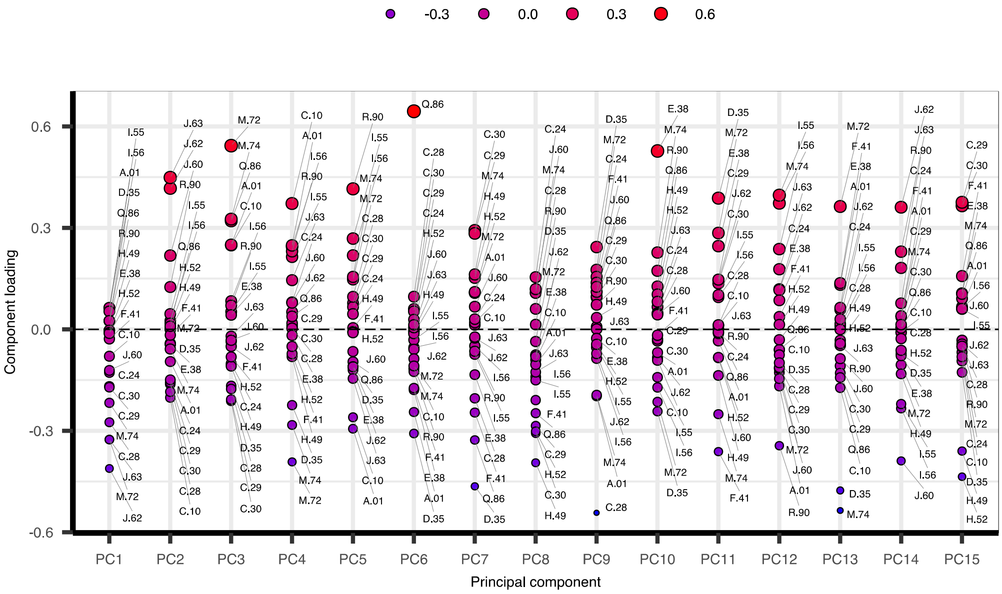


Source: *Own elaboration in R (Blighe and Lun, 2020).*

In the next step, we analyze the component loadings for each industry to describe the specificity of the components (Figure 2). Figure 3 depicts significant correlations between industries and components; hence it is helpful to describe components according to their indication of possible specializations. We observe that although most of the components exhibit distinct industrial domains, some of them are similar in terms of their main core focus. Interestingly, even having the same core industry, we may identify different component orientations, as will be discussed in more detail below.

Interestingly, we do not observe any variables with a positive loading in the case of PC1, hence this component rather indicates what specialization is not. For the most part, this component mirrors the absence of any science or technology and manufacturing industries. It reflects the *'unsophisticated and traditional'* core of regional rural economies with a tenuous focus on conventional tourism, as it is barely fueled by accommodation (I.55), food services (I.56), sport and recreation activities (R.93), creative, arts and entertainment activities (R.90), museums and other cultural activities (R.91) and support services such as travel agencies (N.79), land (H.49) and water (H.50) transport. Inevitably there is an emphasis on agritourism as industries such as agriculture (A.01), forestry (A.02) and fishing (A.03) show modest loadings.

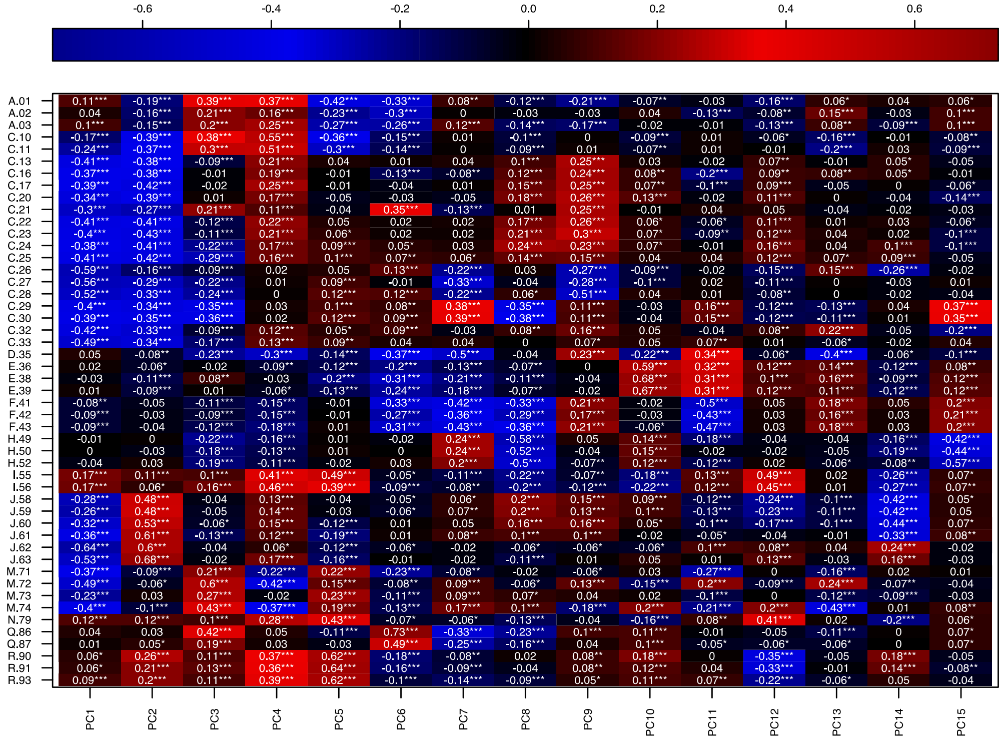
Figure 2. *Component loadings for the chosen principal components and the label industries causing variation along these components*



Source: *Own elaboration in R (Blighe and Lun, 2020). Note: labels correspond to NACE rev. 2 industry codes described in the text.*

There are three more specializations for different types of tourism. PC4 can be named *'experience- and culture-based tourism'* since, in addition to the tourism and catering industries, it refers generally to food products (C.10) and beverages (C.11), as well as crop and livestock production (A.01) and fishing (A.03). Complementary to these activities, this component comprises entertainment, recreation, and amusement activities (R.90, R.91 & R.93). PC5 can be referred to as *'creative industries, culture and tourism'* as the largest loadings originate from: creative, artistic and entertainment activities (R.90), libraries, archives, museums, and other cultural activities (R.91), sports, entertainment, and recreational activities (R.93), and tourism (I.55, I.56 & N.79), excluding food and beverages. This component is also strongly supported by advertising (M.73), scientific research and development (M.72 & M.74) and architectural and engineering activities (M.71). There is another component containing tourism industries (PC12), which can be labeled as *'pure tourism and MICE'* because the loadings derive only from accommodation (I.55), food services (I.56), travel agencies (N.79), and some ancillary services such as bill checking and freight rate information, weather forecasting activities, security consulting (M.74), computer programming (J.62), and information services (J.63).

Figure 3. *Correlation between industries and principal components (specializations) identified in the study*



Source: *Own elaboration in R (Blighe and Lun, 2020). Note: labels correspond to NACE rev. 2 industry codes described in the text.*

PC2 clearly covers *'information and communication technology,'* including information service activities (J.63), computer programming (J.62), telecommunications (J.61), programming and broadcasting activities (J.60), motion picture, video, and television program production, sound recording and music publishing activities (J.59), and publishing activities (J.58). Complementing this major specialization are two others. PC13 integrates research with automated systems, computer, electronic and optical product manufacturing, and construction technologies, so it refers to *'automated manufacturing and construction.'* PC14 indicates a specialization in *'digital arts, entertainment and recreation'* because it combines computer programming (J.62) and information services (J.63) with creative, artistic and entertainment activities (R.90) and libraries, archives, museums, and other cultural activities (R.91).

PC3 comprises research and development (M.72 and 74) in health (Q.86), pharmaceuticals (C.21) and high-quality food and beverages (C.10 and 11), hence it may be termed *'preventive health care'* focused on sustainable quality of life. Another health specialization can be identified in the case of PC6, which can be called *'health and residential care'* as it includes the highest loading of health (Q.86), residential care activities (Q.87) and essential pharmaceuticals (C.21).

PC7 is purely related to the specialty of 'transportation,' which includes both the manufacture of motor vehicles, trailers, and semi-trailers (C.29), other transportation equipment (C.30), and transportation services, including land transportation (H.49), water transportation (H.50), and warehousing and support service activities for transportation (H.52). PC8, on the other hand, appears to be the opposite of 'transportation' specialization, but without specific industries with positive loads. Similarly, PC9 does not show any specific specialty industries with positive loadings, but only indicates a lack of machinery, computer, and electrical equipment. Only the last component analyzed (PC15) indicates specialization in the manufacture of motor vehicles and transport equipment (C.29 and 30), excluding all transport and logistics services.

PC10 addresses the 'environment – sustainability' nexus, pointing to the importance of green technologies. It comprises three industries involved in waste collection, treatment and disposal, materials recovery (E.38), remediation activities and other waste management services (E.39), and water collection, treatment, and supply (E.36), except energy. In turn, PC11 refers to 'clean energy' and technologies for renewable energy production and distribution. This was the last PC of the fifteen components considered.

4. Conclusions

The aim of the paper was to create a typology of smart specializations in European regions based on 246 regional innovation strategies, reported in the Smart Specializations Platform, in which 1,346 specializations were identified. By using PCA analysis, we were able to radically reduce the 82 industries by which specializations were classified. Horn's parallel analysis showed that fifteen MS are appropriate for building a typology of smart specializations. As many as twelve of these fifteen form individual specializations. The identified specializations may constitute the five categories outlined in Table 1.

Table 1. *The typology of smart specializations in Europe*

#	Specialization	PC
	Tourism	
1	Unsophisticated and traditional agritourism	1
2	Experience- and culture-based tourism	4
3	Creative industries, culture, and tourism	5
4	Pure tourism and MICE	12
	ICT	
5	Information and communication technology	2

6	Automated manufacturing and construction	13
7	Digital arts, entertainment, and recreation	14
	Health	
8	Preventive health care	3
9	Health and residential care	6
	Transport	
10	Transportation	7
11	Motor vehicles and transport equipment	15
	Environment	
12	Environment & sustainability	10

Source: *Own elaboration based on PCA results.*

Most specializations in European countries and regions can be classified according to twelve specializations. Interestingly, they all differ from each other, even if they refer to one of the five areas indicated. In other words, they signal a different orientation of the specialization, which is of great importance when analyzing regional strategies, because, for example, agrotourism or cultural tourism differs significantly from MICE and may differently affect regional development.

The only concern that we have noted, and that should be addressed by further research, is that the typology contains a limited number of specializations. That is, the typology includes those specializations (principal components) that significantly explain the variance. And the task of specialization is to differentiate from other regions and if the specialization is unique, it may not be reflected in the adopted typology. This certainly needs to be investigated in future studies.

References:

Barca, F. 2009. An agenda for a reformed Cohesion Policy. A place-based approach to meeting European Union challenges and expectations. Independent report prepared at the request of Danuta Hübner, European Commissioner for Regional Policy.

Blighe, K., Lun, A. 2020. PCAtools: PCAtools: Everything Principal Components Analysis.

Cattell, R.B. 1966. The Scree test for the number of factors. Multivariate Behavioral Research 1, 245-276. https://doi.org/10.1207/s15327906mbr0102_10.

Di Cataldo, M., Monastiriotis, V. 2020. Regional needs, regional targeting and regional growth: an assessment of EU Cohesion Policy in UK regions. Regional Studies 54, 35-47. https://doi.org/10.1080/00343404.2018.1498073.

Foray, D. 2015. Smart specialisation: opportunities and challenges for regional innovation policy. Routledge, Abingdon.

Foray, D. 2014. From smart specialisation to smart specialisation policy. European Journal of Innovation Management 17, 492-507. https://doi.org/10.1108/EJIM-09-2014-0096.

Foray, D., David, P.A., Hall, B. 2009. Smart specialisation – the concept. Knowledge Economists Policy Brief 9, 1-5.

Gavish, M., Donoho, D.L. 2014. The optimal hard threshold for singular values is 4 diveded by root 3. IEEE Transactions on Information Theory 60, 5040-5053. https://doi: 10.1109/TIT.2014.2323359.

Gianelle, C., Guzzo, F., Mieszkowski, K.2020a. Smart Specialisation: what gets lost in translation from concept to practice? null 54, 1377-1388. https://doi.org/10.1080/00343404.2019.1607970.
Gianelle, C., Kyriakou, D., McCann, P., Morgan, K. 2020b. Smart Specialisation on the move: reflections on six years of implementation and prospects for the future. null 54, 1323-1327. https://doi.org/10.1080/00343404.2020.1817364.
Hassink, R., Gong, H. 2019. Six critical questions about smart specialization. European Planning Studies. https://doi.org/10.1080/09654313.2019.1650898.
Horn, J.L. 1965. A rationale and test for the number of factors in factor analysis. Psychometrika 30, 179-185. https://doi.org/10.1007/BF02289447.
Iacobucci, D. 2014. Designing and Implementing a Smart Specialisation Strategy at Regional Level: Some Open Questions. Scienze Regionali 13, 107-126. https://doi.org/10.3280/SCRE2014-001006.
Iacobucci, D., Guzzini, E. 2016. Relatedness and connectivity in technological domains: missing links in S3 design and implementation. European Planning Studies 24, 1511-1526. https://doi.org/10.1080/09654313.2016.1170108.
Jolliffe, I.T. 1986. Principal component analysis, Springer series in statistics. Springer-Verlag, New York.
Kroll, H. 2015. Efforts to Implement Smart Specialization in Practice - Leading Unlike Horses to the Water. null 23, 2079-2098. https://doi.org/10.1080/09654313.2014.1003036.
Mäenpää, A., Teräs, J. 2018. In Search of Domains in Smart Specialisation: Case Study of Three Nordic Regions. EJSD. https://doi.org/10.30689/EJSD2018:68.1650-9544.
Marčenko, V.A., Pastur, L.A. 1967. Distribution of eigenvalues for some sets of random matrices. Mathematics of the USSR-Sbornik 1, 457-483. https://doi.org/10.1070/sm1967v001n04abeh001994.
McCann, P., Ortega-Argilés, R. 2016. The early experience of smart specialization implementation in EU cohesion policy. null 24, 1407-1427. https://doi.org/10.1080/09654313.2016.1166177.
Morgan, K. 2015. Smart Specialisation: Opportunities and Challenges for Regional Innovation Policy. Regional Studies 49, 480-482. https://doi.org/10.1080/00343404.2015.1007572.
Radosevic, S., Ciampi Stancova, K. 2018. Internationalising Smart Specialisation: Assessment and Issues in the Case of EU New Member States. Journal of the Knowledge Economy 9, 263-293. https://doi.org/10.1007/s13132-015-0339-3.
Thorndike, R.L. 1953. Who belongs in the family? Psychometrika 18, 267-276. https://doi.org/10.1007/BF02289263.
Uyarra, E. 2018. Smart specialisation as place-based policy. Lessons learnt? Regions. https://doi.org/10.1080/13673882.2018.00001022.
Vallance, P. 2017. The co-evolution of regional innovation domains and institutional arrangements: smart specialisation through quadruple helix relations? In: McCann, P., van Oort, F., Goddard, J. (Eds.), The Empirical and Institutional Dimensions of Smart Specialisation, Regions and Cities. Routledge, London and New York, 127-144.