UNIMORE
UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

Dipartimento di Economia Marco Biagi

DEMB Working Paper Series

N. 209

Regional innovation in southern Europe: a poset-based analysis

Filippo Damiani[1], Silvia Muzzioli[2], Bernard De Baets[3]

March 2022

[1] University of Modena and Reggio Emilia,
Marco Biagi Foundation
Email: filippo.damiani@unimore.it
[2] University of Modena and Reggio Emilia,
Department of Economics Marco Biagi and CEFIN
Email: silvia.muzzioli@unimore.it
[3] Ghent University, Department of Data Analysis and Mathematical Modelling
Email: bernard.debaets@ugent.be

ISSN: 2281-440X online

Regional innovation in southern Europe: a poset-based analysis

Filippo Damiani,[a] Silvia Muzzioli,[b,*] Bernard De Baets[c]

Abstract

This paper examines the performance of regional innovation across the 60 southern European regions of Greece, Italy, Portugal and Spain. A poset-based analysis is carried out in two phases. The first phase establishes a ranking of the clusters in which regions are grouped to identify patterns of comparable regions. The second phase focuses on the country level, where the regions of each of the four countries are ranked into five different performance levels. The outcomes of the two phases are compared with the results described in the Regional Innovation Scoreboard 2019, with a view to providing insights for policymakers.

Keywords

Poset, European Union, Regional Innovation Scoreboard, Southern Europe, Greece, Italy, Portugal and Spain

The authors wish to thank William Bromwich for his painstaking attention to the editing of this paper.

[a] **filippo.damiani@unimore.it**, *PhD Candidate*
Marco Biagi Department of Economics, University of Modena and Reggio Emilia, Modena, Italy
[b] **silvia.muzzioli@unimore.it**, *Associate Professor*
Marco Biagi Department of Economics, University of Modena and Reggio Emilia, Modena, Italy
[c] **bernard.debaets@ugent.be**, *Senior Full Professor*
KERMIT, Department of Data Analysis and Mathematical Modelling, Ghent University, Ghent, Belgium

[*] Corresponding author: Marco Biagi Department of Economics and CEFIN, University of Modena and Reggio Emilia, Viale Berengario 51, 41121 Modena (I), Tel. +390592056771 Fax +390592056947,ORCID iD: 0000-0003-0738-6690.

1. Introduction

Since the Great Recession of 2008-2009, inequalities among European regions have become more acute, causing a sharp increase in sovereign debt, especially in southern European countries, in particular, Portugal, Italy, Spain, and Greece (Perez & Matsaganis, 2018; Bekiros et al., 2018), where the overall unemployment rate and the youth unemployment rate have increased (Garofalo et al., 2018). The capacity for innovation of these countries has been severely impacted by the crisis, though some of them have been able to innovate even during the worst years of the recession (e.g. Donatiello & Ramella, 2017). Furthermore, the Covid-19 pandemic has made southern European economies even more fragile (Moreira et al., 2020). In this context, the analysis of regional innovation systems has become increasingly important (Rodil-Marzábal & Vence-Deza, 2020) to identify opportunities for economic growth and to secure regional resilience (Coenen et al., 2017).

The literature on regional innovation systems (RISs) has grown significantly in the last decades (Doloreaux & Porto Gomez, 2017). The interest in RISs is driven by the conceptualisation of innovation as a source of competitive advantage (Asheim et al., 2011) and by the linkages between innovation patterns and economic performance (Capello & Lenzi, 2019); arguably, innovation-driven economies can produce new jobs and new value-added products and services (Gabriel, 2019). Several authors consider innovation to be primarily determined at regional level (Doloreux & Parto, 2004; Navarro et al., 2009; Lau & Lo, 2015), while others argue that regions still represent the basic territorial unit for organising the economy (Asheim et al., 2019).

Another relevant aspect related to innovation systems is the importance of cooperation in developing innovative processes to reduce duplication spillovers while sharing costs and risks (Nunes et al., 2013). Collaboration is widely considered a key element also to achieve higher regional innovation performance and achieve greater innovation synergy effects (Ponsiglione et al., 2018; Russell & Smorodinskaya, 2018). Moreover, interregional linkages have a positive effect on the likelihood of regions to diversify, especially for regions that showcase complementary capabilities (Balland & Boschma, 2021). For this reason, identifying peer regions is of crucial importance to create a powerful learning channel and positively influence innovation policies (Franco et al., 2020).

Among the different indices available in the literature to measure regional innovation performance, one of the most widely adopted is the Regional Innovation Scoreboard, which has been adopted by several scholars (Zabala-Iturriagagoitia et al., 2007; Arbolino et al., 2019; Garcia-Bernabeu et al., 2020), some of whom consider it to be the most important innovation index at the regional level (Hauser et al., 2018). The ninth Regional Innovation Scoreboard (RIS, for short) was published in 2019 and provides a comparative assessment of the performance of innovation systems across 238 European regions (Hollanders et al., 2019a); the final score of the RIS is calculated as the unweighted average of the normalised scores of 17 indicators, which constitutes a severe limitation since such aggregation is subject to possible compensation effects (Carlsen, 2018). An attempt to revisit the RIS has been proposed, for instance, by Carayannis et al., 2018. In particular, these authors employ a Multiple-Criteria Decision Analysis (MCDA) approach combining the AHP and the TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution) methods.

In this paper, we adopt an approach borrowed from the theory of partially ordered set (theory of posets, or poset theory, for short). The poset-based approach can be used as an alternative to composite indicators and facilitates the process of ranking in an insightful manner (Brüggemann & Patil, 2011). Fuelled by powerful algorithms (De Loof et al., 2008), it has been adopted in a wide range of studies in the literature, including the calculation of new indices of stringency of fiscal rules (Badinger & Reuter, 2015), the synthetisation of multi-indicator systems over time (Alaimo et al., 2020), the evaluation of multidimensional poverty (Fattore & Arcagni, 2014), the quality assessment of river water (Tsakovski et al., 2010), and the statistical evaluation of socio-economic phenomena (Fattore et al., 2012). The main strengths of the poset-based approach can be summarised as follows: it respects the ordinal nature of data, it maintains a high standard of objectivity (reducing the need for subjective choices), and it fully exploits all the information in the dataset (Badinger & Reuter, 2015). We believe that the aforementioned characteristics enable it to capture and represent the complexity of the measurement of regional innovation performance. Through these characteristics, it is possible to identify the most impacting indicators and consider them as all relevant in the construction of the ranking.

Our aim is to apply the poset-based approach to the RIS 2019 data of the 60 regions of southern Europe to identify similarities and differences with regard to their innovation performance. The findings of our analysis are intended to be of interest for policymakers considering that regional authorities need to tailor their own place-based policies (Grillitsch & Asheim, 2018; Morrison & Doussineau, 2019) as one-size-fits-all policies are not the solution, especially for regions lagging behind (Rodríguez-Pose & Ketterer, 2020). However, grouping similar regions into clusters can help connect policies to tackle challenges in a more focussed manner (Mazzucato, 2018). The analysis consists of three steps: first, a clustering of the 60 regions of the four countries (Greece, Italy, Portugal and Spain) is performed; second, the poset-based approach is applied to establish a ranking of these clusters (first phase); third, the first two steps are repeated considering only the data of the individual countries to find even more detailed evidence (second phase). The application of the poset-based approach is feasible even on a fairly large dataset, thanks to the cluster analysis and the attribute-related sensitivity analysis proposed in this paper. This enables us to identify the indicators with the greatest impact for each of the countries analysed. We finally compare our results with the ones described in the RIS 2019.

This paper is organised as follows. Section 2 presents the dataset and the methods used, with particular attention to the description of the various steps of the analysis. Section 3 presents the results of the study. The last section is dedicated to the discussion of findings, conclusions, limitations, and perspectives for future research. Appendix A outlines the Regional Innovation Scoreboard 2019. Appendix B provides a data analysis example using the poset-based approach. Appendix C describes the 60 regions analysed in this study, listed in alphabetical order. It also provides additional information about the clusters and the assignment of the regions to a specific performance level resulting from the two phases of our analysis and from the RIS 2019 report. This description is intended to promote a better understanding of the regions that are similar in terms of innovation performance, and a clearer idea of which indicators should be prioritised to improve the position of a region in the ranking thanks to the attribute-related sensitivity analysis.

2. Material and methods

In this section, we describe the dataset adopted for the study of regional innovation in southern Europe and the methods adopted in the different steps of our investigation. The analysis was carried out adopting the dataset downloaded from the Regional Innovation Scoreboard 2019 website (https://bit.ly/3cc8PAP). The scores are already normalised for all indicators. We exclude the indicator *SMEs non-R&D innovation expenditures as percentage of GDP* from the dataset for the reasons explained in Appendix A. As a result, in our analysis we consider the 60 regions (NUTS[1] 2) of Greece, Italy, Portugal and Spain and 16 indicators divided into four different frameworks. Only 21 values are missing, most of which (13 out of 21) belong to the indicator *employment in medium-high and high-tech manufacturing and knowledge-intensive services as percentage of total employment*.

The first step of the analysis is the imputation of the 21 missing values. To this end, we adopted the nearest neighbour imputation method, a commonly applied method (Jadhav et al., 2019). More precisely, we considered the five nearest neighbour values to compute each item of missing data. The imputation was carried out for each indicator separately. After imputation, the data matrix contains 960 observations.

The application of the poset-based approach to a large dataset could generate results that are difficult to interpret. With this in mind, we reduced the number of objects (60 regions) through a cluster analysis by performing a hierarchical clustering with the default distance measure, namely the Euclidean distance measure; the function used is "hclust" with the complete linkage method (using the software R). The scores of the clusters correspond (for each attribute) to the average of the scores of the regions that compose each cluster. The number of attributes (indicators) is then reduced to two, for each of the four categories, through the attribute-related sensitivity method (see Appendix B, Table B.3). After the reduction of both the number of objects and attributes, the poset-based approach is applied to the final

[1] NUTS stands for the Nomenclature of Territorial Units for Statistics and it is used for referencing the subdivisions of countries.

data matrix to create a ranking of the clusters of the 60 southern European regions (first phase of the analysis). The same procedure is then applied considering just one country at a time to detect even more differences among the regions of the same country (second phase of the analysis). This country-level focus can be considered as a robustness analysis to validate the ranking obtained in the first phase.

In the last step of the analysis, we provide a comparison between the performance levels of southern European regions obtained in the poset-based analysis and the performance groups described in the RIS 2019. Finally, we provide a comparison of the indicators with the greatest impact resulting from the attribute-related sensitivity analysis, which is carried out five times (first it is applied to the attributes of all 60 regions, then to the attributes of just the regions of the individual countries). We illustrate the characteristics of the poset-based approach in Appendix B.

3. Results

In this section, we describe the results of the analysis conducted both at regional level for the 60 southern European regions, and at country level for Greece, Italy, Portugal and Spain. Subsections 3.1 and 3.2 contain the results for the cluster and the attribute-related sensitivity analyses. Subsections 3.3 and 3.4 contain the results of the poset-based analysis at regional and country level, respectively.

3.1 Cluster analysis

As explained in the previous section, after the imputation of the missing values, the dataset consists of the 60 regions of southern European countries and includes 16 indicators. As this is too large a dataset to be analysed with the basic poset-based approach, it is necessary to create clusters of regions.

The first step is the computation of the distance matrix, showing for each pair of objects (regions) their Euclidean distance considering all the indicators. The clusters are then created based on the distance matrix according to the complete linkage method.

We chose a number of clusters (*k*) equal to nine to be able to reduce the ‘within group sum of squares’ and at the same time obtain a sufficiently rich partial order, by which we can identify five different performance levels after the data analysis. As a result, the matrix consists of nine rows (clusters of regions) and 16 columns (indicators). The scores of a cluster are the averages of the scores, between 0 and 1 (normalised values), of the regions that compose the cluster.

According to the cluster analysis performed, the number of regions included in the different clusters is not homogeneous. In particular, we observe, on the one hand, larger clusters consisting of ten (clusters 4 and 8) or more regions (cluster 9), and on the other hand, one cluster that includes just one region (cluster 7).[2] Hence, the first result is that the Spanish autonomous city of Ceuta (located on the coast of north Africa) shows data that are incomparable with all other regions included in the dataset, and, with $k = 9$, it is impossible to include it in any cluster.

3.2 Attribute-related sensitivity analysis

In this step of the analysis, we aim to select the two indicators with the greatest impact, that is, the two most impacting factors for each of the four categories to reduce the number of indicators from 16 to 8. Since the two categories 'investments' and 'impacts' consist of two indicators each, it is not necessary to perform any reduction for them. As a result, we apply the attribute-related sensitivity analysis to the two remaining categories. We reduce the four indicators of the category 'framework conditions' and the eight indicators of the category 'innovation activities'.

Starting with 'framework conditions', we consider a data matrix consisting of the nine clusters as objects and the four indicators of the category under analysis. After obtaining the Hasse diagram representing the relations between the clusters for this category, it is important to compute the total number of incomparabilities as an estimate of the complexity of the poset, and then find the pair of indicators that reproduces the closest number of incomparabilities. There are 22 incomparabilities in the Hasse diagram generated considering all four attributes of the category. The indicators *percentage of population aged 30-34 with tertiary education* and *top-10% most cited publications worldwide as percentage of total scientific publications of the country* alone create 15 incomparabilities (68% of the total); thus, since they are the ones with the greatest impact for the category, they will be considered in the final data matrix.

Regarding the category 'Innovation activities', there are eight indicators. As a result, the number of possible pairwise combinations is quite high. In this case, there are 31 incomparabilities and the pair of indicators with the strongest impact is formed by *SMEs introducing marketing or organisational innovations as percentage of SMEs* and *European design applications per billion GDP in PPS*, representing 24 incomparabilities (77% of the total).

The final data matrix in Table 1 shows the nine clusters and eight indicators, representing the two with the strongest impact for each category and listed as follows (the indicators are identified as indicated in Appendix A: 1. Framework conditions: *1a. percentage of population aged 30-34 with tertiary education*; *1d. top-10% most cited publications worldwide as percentage of total scientific publications of the country*. 2. Investments: *2a. R&D expenditure in public sector as percentage of GDP*; *2b. R&D*

[2] As a number of clusters contain several regions, the variability inside these clusters is likely to be quite high. As a result, some regions might be considered as outliers of such clusters, as in the case of Dytiki Makedonia (cluster 1), Kriti (cluster 2), Notio Aigaio (cluster 3), and Região Autónoma da Madeira and Valle d'Aosta (cluster 9).

expenditure in business sector as percentage of GDP. 3. Innovation activities: *3b. SMEs introducing marketing or organisational innovations as percentage of SMEs*; *3h. European design applications per billion GDP in PPS*. 4. Impacts: *4a. Employment in medium-high and high-tech manufacturing and knowledge-intensive services*; *4b. SMEs sales of new-to-market and new-to-firm innovations as percentage of total turnover*.

Table 1 - Final data matrix: nine clusters and eight indicators with the greatest impact (data normalised)

Cluster	*1a*	*1d*	*2a*	*2b*	*3b*	*3h*	*4a*	*4b*
1	0.349	0.356	0.389	0.140	0.544	0.141	0.184	0.669
2	0.492	0.440	0.601	0.274	0.664	0.247	0.283	0.683
3	0.339	0.335	0.418	0.102	0.602	0.118	0.158	0.342
4	0.469	0.376	0.407	0.241	0.300	0.206	0.271	0.556
5	0.650	0.486	0.513	0.523	0.371	0.316	0.641	0.667
6	0.440	0.421	0.438	0.256	0.313	0.570	0.257	0.547
7	0.226	0.000	0.104	0.007	0.105	0.388	0.215	0.384
8	0.285	0.505	0.491	0.465	0.511	0.697	0.505	0.666
9	0.188	0.520	0.446	0.278	0.463	0.251	0.354	0.631

Source: elaborated on the results obtained from software R.

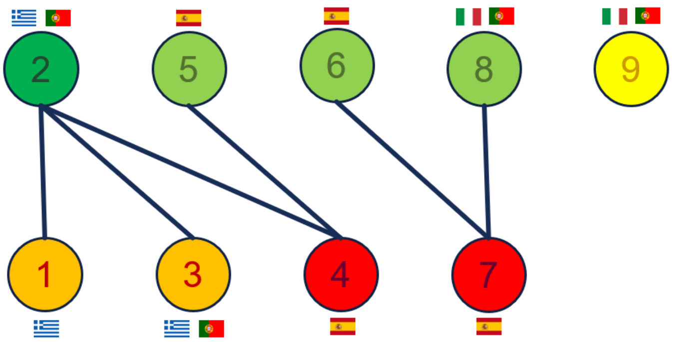
The entries for cluster 7 are just those for Ceuta since it is the only region in this *sui generis* cluster.

The results of the first phase of the analysis (the ranking of the clusters considering all 60 southern European regions) are presented in Subsection 3.3, whereas the results of the second phase (country-level analysis) are reported in Subsections 3.3.1 to 3.3.4.

3.3 Results of the poset-based analysis considering all 60 southern European regions

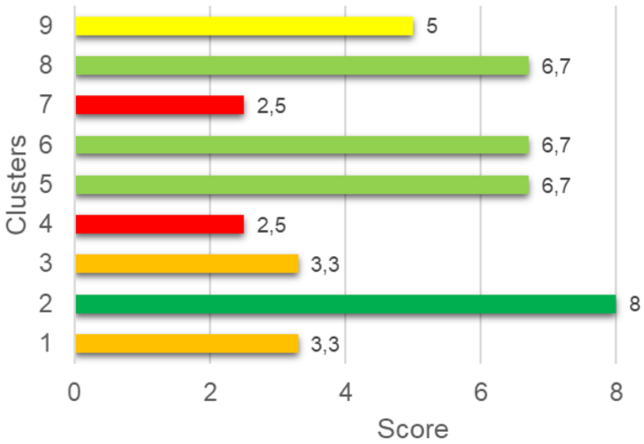
The Hasse diagram obtained from the data matrix in Table 1 is shown in Figure 1.

Figure 1 – Nine clusters (60 southern European regions), Hasse Diagram



The Hasse diagram clearly shows the relations between the clusters. In examining Figure 1, we can divide the clusters into five different levels according to the relations between them. Cluster 2 (consisting of regions from Greece and Portugal) is better than three different clusters (1, 3, and 4), and we expect it to be the first cluster in the ranking; clusters 5, 6, and 8 are better than just one cluster each (cluster 5 is better than cluster 4, and clusters 6 and 8 are both better than cluster 7); cluster 9 is incomparable with all the other clusters; clusters 1 and 3 are worse than just one cluster, which is cluster 2 in both cases; finally, clusters 4 and 7 (Spanish regions only) are worse than two clusters each (cluster 4 is worse than clusters 2 and 5, and cluster 7 is worse than clusters 6 and 8). The ranking in Figure 2 is obtained by applying the Local Partial Order Model (LPOM).

Figure 2 – Final scores of the clusters obtained by applying the LPOM



The Local Partial Order Model highlights five levels of performance: the top level, consisting of cluster 2; the middle-top level consisting of clusters 5, 6 and 8 (all of them with the same score); the middle level, consisting of cluster 9 (the cluster that is incomparable with all the others); the middle-bottom level, consisting of clusters 1 and 3 (with the same score); finally, the bottom level, consisting of clusters 4 and 7, again with the same score. More detailed results are provided in Figure 3, which shows the composition of each cluster and gives information about the number of regions for each country.

Figure 3 – Composition of the nine clusters (60 southern European regions)

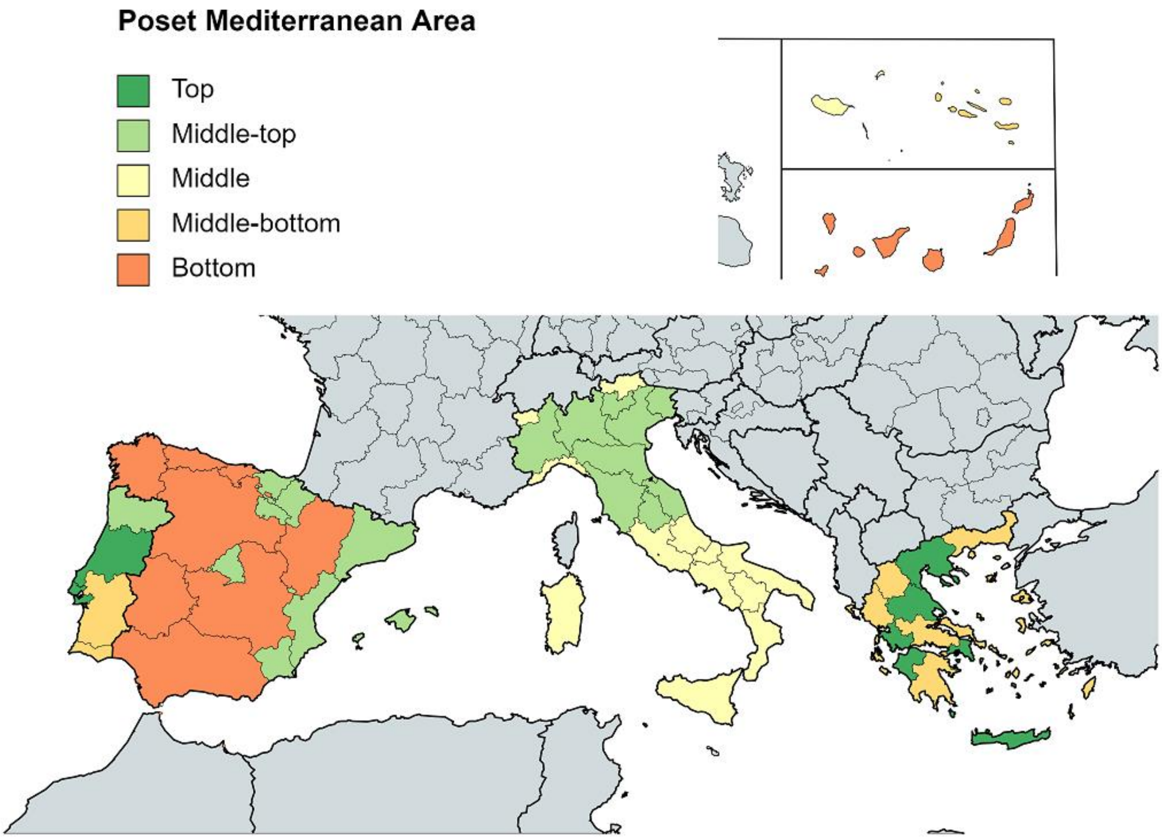
Cluster	EL	ES	IT	PT	TOTAL
2	5			2	7
5		4			4
6		4			4
8			9	1	10
9			12	1	13
1	5				5
3	3			3	6
4		10			10
7		1			1
TOTAL	13	19	21	7	60

	EL	ES	IT	PT	TOTAL
Top	5	0	0	2	7
Mid-Top	0	8	9	1	18
Middle	0	0	12	1	13
Mid-Bottom	8	0	0	3	11
Bottom	0	11	0	0	11

The left-hand panel in Figure 3 orders the clusters from top performing to bottom performing. As regards the colours, dark green represents the top-performing cluster (just cluster 2), light green represents the middle-top performing ones, yellow stands for the cluster in the middle level (cluster 9), orange represents the middle-bottom performing clusters, and red the bottom performing ones. The rows indicate the number of regions of each country that compose the clusters, whereas the columns indicate the clusters into which the regions of the different countries are divided. The right-hand panel of Figure 3 shows the number of regions (country by country) that compose each of the five levels: the number of regions is almost evenly distributed over the performance levels, with 25 regions placed in the first two levels, 22 regions forming the last two levels, and 13 in the middle. The top level is composed of five regions from Greece and two from Portugal, whereas the bottom level is composed only of Spanish regions. Moreover, Italian regions are the only ones that are neither in the top level, nor in the bottom level; in fact, nine Italian regions are in the middle-top level and the other 12 are in the middle level.

The results of the 60 regions analysed are displayed on a political map in Figure 4.

Figure 4 – Political map representing the result of the poset-based analysis of 60 southern European regions



An examination of Figure 4 provides more insights. In Italy we observe a clear difference between the north and the south: the northern regions belong to the middle-top level (except for Valle d'Aosta, Provincia Autonoma di Bolzano, and Liguria), whereas the southern regions pertain to the middle level. The same pattern can be observed in Portugal: the northern regions are in the top or middle-top level, whereas the southern regions are in the middle-bottom level. More heterogeneous performances are visible in Spain and in Greece: in the case of Spain, it is evident that the best performing regions are located in the north-east and south-east of the country (including Madrid); in Greece, the top regions are equally distributed over the national territory. Last but not least, none of the regions in which the capital city is located belongs to a bottom (or middle-bottom) level cluster: Lisbon and Athens are both top-level regions, Madrid is a middle-top one, and Rome belongs to a middle-level region, namely Lazio (even if Rome is located in a large region compared to the aforementioned ones, which can concentrate all their resources in high-density areas).

If we intend to try to find more details about the differences in performance of regions of the same country, it is necessary to repeat the analysis considering just one country at a time. To do so, we performed a cluster analysis for each country to identify the number of clusters enabling us to obtain a sufficiently rich partial order to identify five performance levels for each country, as in the first phase of the analysis. In some cases, we have more clusters than performance levels (as in the first phase) since two or more clusters could have the same score, meaning that they will be assigned to the same performance level. At the same time, the attribute-related sensitivity analysis is also carried out for each country to find the indicators with the greatest impact. An interesting comparison of the different analyses to identify the most impacting indicators is presented in subsection 3.4. The country-level focus (presented in the following subsections) also serves as a robustness analysis.[3] In fact, the aim of the second phase of the analysis is not only to obtain more detailed results for each country, but also to check the consistency of the rankings obtained in the second phase with the ones obtained in the first phase, which means that, despite changes in the indicators with the greatest impact and the clusters in which regions are grouped, the order of regions in the ranking is not inverted.

3.3.1 A focus on Greece

The 13 regions of Greece were placed in three different clusters in the first phase of the analysis distributed in just two performance levels: five regions in the top level, and eight regions in the middle-bottom one. To obtain five performance levels the only possibility is to group the Greek regions into six clusters in the country-level analysis. The largest cluster consists of four regions (Attiki, Dytiki Ellada, Kentriki Makedonia, and Thessalia), whereas there is one cluster consisting of just one region (Kriti). The results of the second phase of the analysis for Greece are shown in Table 2, along with the results for the same regions in the first phase.

[3] The robustness analysis has to be seen as a check of the order of the regions obtained in the first phase compared with the order of the same regions in the second phase. The comparison between the level of a given region in the two phases is not significant since the levels are constructed considering different scales in the two phases (60 regions in the first phase and just the regions of the country analysed in the second phase).

Table 2 - Results for Greek regions in the two phases of the analysis

Region	First phase (level)	Second phase (level)
Kriti	Top	Top
Attiki	Top	Middle-top
Kentriki Makedonia	Top	Middle-top
Dytiki Ellada	Top	Middle-top
Thessalia	Top	Middle-top
Dytiki Makedonia	Middle-bottom	Middle
Ipeiros	Middle-bottom	Middle
Voreio Aigaio	Middle-bottom	Middle
Notio Aigaio	Middle-bottom	Middle-bottom
Sterea Ellada	Middle-bottom	Bottom
Ionia Nisia	Middle-bottom	Bottom
Anatoliki Makedonia, Thraki	Middle-bottom	Bottom
Peloponnisis	Middle-bottom	Bottom

Table 2 clearly shows the differences for the Greek regions between the first and the second phase. Comparing the two phases of the analysis, it is clear that the five regions that were placed in the same top-level cluster in the first phase are now divided into two clusters: Kriti is better ranked than the other four regions. Regarding the eight regions that in the first phase were at the middle-bottom level, they are now divided into three different performance levels in the country-level analysis: Dytiki Makedonia, Ipeiros, and Voreio Aigaio (middle level) are better ranked than Notio Aiagio (middle-bottom level), Sterea Ellada, Ionia Nisia, Anatoloki Makedonia-Thraki, and Peloponnisis (bottom level). The results of the country-level analysis show for all regions of Greece that the ranking is consistent with the ranking obtained in the first phase, and more details about the relations between regions are found.

3.3.2 A focus on Italy

To obtain five different levels of performance of the 21 Italian regions, it is necessary to divide them into eight clusters. The largest cluster consists of five regions (Abruzzo, Basilicata, Campania, Molise, and Puglia), whereas two clusters consist of just one region each (Bolzano and Valle d'Aosta), which are incomparable with all other clusters. The results of the second phase of the analysis for Italy are shown in Table 3, as well as the results for the same regions in the first phase.

Table 3 - Results of Italian regions in the two phases of the analysis

Region	First phase (level)	Second phase (level)
Friuli-Venezia Giulia	Middle-top	Top
Provincia Autonoma Trento	Middle-top	Top
Toscana	Middle-top	Top
Lazio	Middle	Middle-top
Liguria	Middle	Middle-top
Emilia-Romagna	Middle-top	Middle
Lombardia	Middle-top	Middle
Marche	Middle-top	Middle
Piemonte	Middle-top	Middle
Umbria	Middle-top	Middle
Veneto	Middle-top	Middle
Provincia Autonoma Bolzano	Middle	Middle
Valle d'Aosta	Middle	Middle
Abruzzo	Middle	Middle-bottom
Basilicata	Middle	Middle-bottom
Campania	Middle	Middle-bottom
Molise	Middle	Middle-bottom
Puglia	Middle	Middle-bottom
Calabria	Middle	Bottom
Sardegna	Middle	Bottom
Sicilia	Middle	Bottom

In the first phase of the analysis, the Italian regions were placed in just two clusters: nine regions were in a cluster at the middle-top level and the other 12 regions in a cluster positioned at the middle level. The second phase of the analysis reveals more details about the differences in the performance of the Italian regions. On the one hand, the nine regions in the middle-top cluster of the first phase of the analysis are now divided into two different performance levels: Friuli-Venezia Giulia, Trento, and Toscana constitute the top level of the analysis at the country level, whereas Emilia-Romagna, Lombardia, Marche, Piemonte, Umbria, and Veneto are placed in the middle level. On the other hand, the 12 regions belonging to the middle level in the first phase are now divided into three different performance levels. In fact, Lazio and Liguria are now in the middle-top level, whereas the eight regions of the south are in two different clusters: Abruzzo, Basilicata, Campania, Molise, and Puglia are in a middle-bottom level cluster, and are ranked in a better position than Calabria, Sardegna, and Sicilia, which are placed in a bottom-level performance cluster. The results of the country-level analysis for the Italian regions demonstrate that the ranking obtained in the first phase is consistent for 19 regions out of 21. Liguria and Lazio were at a lower level than nine regions in the first phase of the analysis, whereas in the second phase they are ranked in a lower position than only three other regions, surpassing the regions of Emilia-Romagna, Lombardia, Piemonte (together these three regions form a cluster in the country-level analysis), Marche, Umbria, and Veneto (another cluster of three regions in the second phase of the analysis). These last two clusters are strongly penalised in the second phase due to a modest performance for just one indicator, namely *R&D expenditure in public sector as percentage of GDP*, which makes these clusters incomparable with all the others, even if for all the other indicators these two clusters are better performing than most of the other Italian clusters. This confirms that in the poset-based analysis it is not sufficient to obtain a 'good mean' score, but that it is fundamental not to have a low performance score on any indicator to avoid being downgraded in the ranking.

3.3.3 A focus on Portugal

Since Portugal has just seven regions, it is not necessary to conduct a cluster analysis to obtain five different performance levels. Table 4 shows the results of the analysis of the second phase (regarding Portugal), compared with the results of the same regions according to the analysis conducted in the first phase.

Table 4 – Results of Portuguese regions in the two phases of the analysis

Region	First phase (level)	Second phase (level)
Lisboa	Top	Top
Centro	Top	Middle-top
Norte	Middle-top	Middle
Região Autónoma da Madeira	Middle	Middle
Algarve	Middle-bottom	Middle
Região Autónoma dos Açores	Middle-bottom	Middle-bottom
Alentejo	Middle-bottom	Bottom

In the first phase of the analysis, the regions of Portugal belonged to four different clusters (as shown in Figure 3) positioned on four different levels of the ranking (from top to middle-bottom). In the analysis with all countries, Lisbon and Centro were in the same top-level cluster. However, as shown in the last column of Table 4, in the country-level analysis they are divided into two different performance levels: Lisbon is still at the top level, whereas Centro is now at a middle-top level. Furthermore, in the second phase, the regions of Algarve, Madeira, and Norte are incomparable among themselves and with all the other Portuguese regions; hence, they constitute the middle level. Furthermore, Algarve, together with the region of the Açores and Alentejo, were all placed in the same middle-bottom level cluster in the analysis with all countries, and it was not possible to establish a clear relation between them. Thanks to the country-level analysis, it may be seen that Algarve is better ranked than the region of the Açores, which is better ranked than Alentejo. The results of the second phase, concerning Portuguese regions, show that the ranking is consistent with the ranking obtained in the first phase. Moreover, we are now able to identify relations among those regions that in the first phase were placed in the same clusters. However, if we consider the results of the second phase, it seems that Norte, Madeira, and Algarve are at the same performance level. In this case, the combination between the first and the second phases could be helpful in understanding the ranking among these regions.

3.3.4 A focus on Spain

The 19 regions of Spain were placed in four different clusters in the first phase of the analysis. To obtain five performance levels, it is sufficient to divide the Spanish regions into five clusters in the country-level analysis. This means that the clusters show a clear pattern considering the indicators with the greatest impact: each cluster corresponds to one performance level. The results of the second phase of the analysis for Spain are shown in Table 5, along with the results for the same regions in the first phase.

Table 5 - Results for the Spanish regions in the two phases of the analysis

Region	First phase (level)	Second phase (level)
Cataluña	Middle-top	Top
Comunidad de Madrid	Middle-top	Top
Comunidad Foral de Navarra	Middle-top	Top
País Vasco	Middle-top	Top
Comunidad Valenciana	Middle-top	Middle-top
Islas Baleares	Middle-top	Middle-top
La Rioja	Middle-top	Middle-top
Murcia	Middle-top	Middle-top
Aragón	Bottom	Middle
Asturias	Bottom	Middle
Cantabria	Bottom	Middle
Galicia	Bottom	Middle
Andalucía	Bottom	Middle-bottom
Canarias	Bottom	Middle-bottom
Castilla-la Mancha	Bottom	Middle-bottom
Castilla y León	Bottom	Middle-bottom
Extremadura	Bottom	Middle-bottom
Ciudad Autónoma de Melilla	Bottom	Middle-bottom
Ciudad Autónoma de Ceuta	Bottom	Bottom

As in the case of the Greek regions, for the Spanish regions the differences between the first and the second phase are evident, as can be seen in Table 5. In the first phase of the analysis the 19 regions were grouped into four different clusters belonging to just two performance levels: eight regions in the middle-top level, and 11 regions in the bottom level. In the second phase more insights are available. The eight regions in the middle-top level in the first phase of the analysis are now divided into two different groups: Cataluña, Madrid, Navarra, and País Vasco compose the top level in Spain and are ranked higher than Comunidad Valenciana, Islas Baleares, La Rioja, and Murcia, which form the middle-top level in the second phase. Furthermore, the 11 regions grouped in the bottom level in the first phase, are divided into three different performance levels in the country-level analysis: Aragón, Asturias, Cantabria, and Galicia (middle level) rank higher than Andalucía, Canarias, Castilla-la Mancha, Castilla y León, Extremadura, Melilla (middle-bottom level), and Ceuta (bottom level). The results of the country-level analysis also show that for Spain the ranking is consistent with the ranking obtained in the first phase. Moreover, in line with the results for Greece, Italy, and Portugal, for the Spanish regions we are able to identify relations between those regions that in the first phase were placed in the same clusters or performance level.

3.4 Indicators with the greatest impact

During the analysis conducted in this paper in the application of the poset-based approach we adopted many times the attribute-related sensitivity analysis to reduce the number of indicators of the first category (framework conditions) and the third one (innovation activities). In Tables 6 and 7 we show respectively the indicators identified as having the greatest impact in the different analyses for the two categories mentioned above. Table 6 represents the indicators of the first category and Table 7 the indicators of the second category; 1 signifies that the indicator has the greatest impact, otherwise it is 0. The last row of Tables 6 and 7 shows the number of cases in which the indicator has the greatest impact.

Table 6 – Indicators with the greatest impact resulting from the attribute-related sensitivity analysis for the category 'framework conditions'

Type of analysis	*1a*	*1b*	*1c*	*1d*
All 60 regions (first phase)	1	0	0	1
Greece	1	0	0	1
Italy	1	0	1	1
Portugal	0	1	0	1
Spain	0	1	0	0
Importance of the indicators[4]	**3**	**2**	**1**	**4**

The indicators with the greatest impact resulting from all the analyses conducted for the first category are *1a. percentage of population aged 30-34 with tertiary education,* and *1d. top-10% most cited*

[4] The importance is measured as the number of times the indicator is among the most impacting ones in the five analyses. Range of the impact: 0-5.

publications worldwide as percentage of total scientific publications of the country, as shown in Table 6. The number of indicators with the greatest impact resulting from the attribute-related sensitivity analysis could be higher (or lower) than two, as in the case of Italy (three indicators with the greatest impact) and Spain (one). In the case of Italy there are three indicators with the greatest impact since the comparison of two pairs of indicators (with only three indicators in the two pairs) resulted in an *ex aequo*. In the case of Spain there is just one impacting indicator. Indicator *1d* is most impacting in all the analyses, except for the Spanish analysis at the country-level. Spain and Portugal are the only two countries for which the indicator *lifelong learning* has a greater impact. Finally, the indicator *international scientific co-publications per million population* is most impacting for Italy only.

Table 7 – Indicators with the greatest impact resulting from the attribute-related sensitivity analysis for the category 'innovation activities'

Type of analysis	*3a*	*3b*	*3c*	*3d*	*3e*	*3f*	*3g*	*3h*
All 60 regions (first phase)	0	1	0	0	0	0	0	1
Greece	0	1	0	0	0	0	1	1
Italy	0	1	0	0	1	0	1	1
Portugal	0	1	1	0	0	0	1	0
Spain	0	0	0	0	1	0	0	1
Importance of the indicators	**0**	**4**	**1**	**0**	**2**	**0**	**3**	**4**

Table 7 shows that the most impacting indicators resulting from all the analyses conducted for the third category are *3b. SMEs introducing marketing or organisational innovations as percentage of SMEs*, and *3h. European design applications per billion GDP in PPS*. Indicator *3g. trademark applications per billion GDP in PPS* also resulted often as most impacting. In this category Italy has four indicators that are all most impacting, Portugal and Greece have three, and Spain just two, as well as two that are the most impacting indicators resulting from the first phase of the analysis with all the 60 southern European regions. There are two more indicators that, in some cases, are among the most impacting ones: indicator *3e. public-private co-publications per million population* (for Italy and Spain), and indicator *3c. SMEs innovating in-house as percentage of SMEs* (for Portugal).

In the next, and last, subsection we provide a comparison to highlight the similarities and differences between the results of the data analysis of the poset-based approach conducted in this study with the one described in the Regional Innovation Scoreboard 2019 report.

3.5 Poset-based analysis vs RIS 2019: a comparison between the two rankings

We now compare the position of each region in the five performance levels of the first phase of our analysis with the position of the same region in the five performance groups of the RIS 2019, bearing in mind that the top, middle-top, middle, middle-bottom, and bottom levels of the poset-based analysis are comparable respectively to the strong, moderate+, moderate, moderate-, and modest groups indicated in the RIS 2019 report. The results are shown in Table 8 below.

Table 8 - A comparison between the composition of the performance groups in the RIS 2019 and in the first phase of the poset-based analysis

Performance groups	**1. Poset Top**	**2. Poset Middle-Top**	**3. Poset Middle**	**4. Poset Middle-Bottom**	**5. Poset Bottom**
1. RIS Strong	**3**	**2**	0	0	0
2. RIS Moderate+	**2**	**10**	**0**	0	0
3. RIS Moderate	2	**5**	**8**	**7**	2
4. RIS Moderate-	0	1	**5**	**3**	**4**
5. RIS Modest	0	0	0	**1**	**5**

Table 8 presents some significant findings. First, 29 out of 60 regions are in the same performance level both in the RIS 2019 and in the poset-based analysis. Second, 26 regions are in adjacent performance levels in both analyses. Third, all Strong Innovator and Moderate+ Innovator regions in the RIS 2019 analysis are in the top level or in the middle-top level in the poset-based analysis. Fourth, all Modest Innovator regions in the RIS 2019 are in the middle-bottom or the bottom level in our analysis. Finally, just five regions are two levels apart in the two analyses. These five regions are: Dytiki Ellada and Thessalia, two Greek regions that are classified as Moderate Innovators in the RIS 2019 but appear in the top level in our analysis; Aragón and Cantabria, two Spanish regions that are in the Moderate Innovators group in the RIS 2019, but appear in the bottom level in our analysis; finally, Isles Baleares (Spain), considered to be a Moderate-Innovator in the RIS 2019, appears in the middle-top level in our analysis.

The regions of Dytiki Ellada and Thessalia have the lowest scores of cluster 2, which is the highest ranked in the poset-based analysis. However, they are not outliers of the cluster, and their positive results are confirmed also in the country-level analysis, in which they have lower performance scores than just one Greek region (Kriti). Furthermore, they perform well on nearly all of the most impacting indicators. Hence, with a cluster analysis performed to obtain five performance levels in the poset-based analysis, Dytiki Ellada and Thessalia should be considered positive performers in terms of innovation.

Aragón and Cantabria have the highest scores of cluster 4, which is the lowest ranked (together with cluster 7) in the poset-based analysis. However, as in the case of the two Greek regions discussed above, they cannot be considered as outliers of their cluster, even if their scores are better compared to the average of the regions of the cluster. In fact, in the country-level analysis, the 10 regions of cluster 4 of the first phase are divided into three clusters, and Aragón and Cantabria (together with Asturias and Galicia) rank higher than the other six regions of cluster 4. Hence, in our analysis, Aragón and Cantabria should be included in a bottom-level cluster, but thanks to the country-level analysis we can conclude that they rank higher than the other regions of the cluster.

Lastly, Isles Baleares is included in a cluster that performs well on the indicators with the greatest impact, both in the first phase and in the country-level analysis, even if the mean of all indicators is not particularly high for the region, nor for the cluster. In particular, Isles Baleares (as well as its cluster) has one of the highest scores for the indicator *European design applications per billion GDP in PPS*, which is one with the greatest impact of the whole analysis.

4. Discussion and conclusions

The aim of this study was to provide a detailed analysis of the innovation performance of the 60 regions of southern Europe, providing more insights than the simple ranking based on just the unweighted average of the normalised scores, as suggested by the Regional Innovation Scoreboard. To achieve our goal, we analysed the data of the RIS 2019, which is the most recent version available of the Scoreboard, in two phases. In the first phase, we considered all the 60 regions together. In the second phase, we carried out four different country-level analyses, one for each country. The two phases consist of the same steps: the division of regions into clusters, the identification of the two most impacting indicators for each of the four categories of indicators, the computation of the ranking of the clusters of regions in order to identify five levels of performance (top, middle-top, middle, middle-bottom, and bottom level).

In the first phase, we divided the 60 southern European regions into nine clusters according to the similarity of their scores for the indicators. The information resulting from the cluster analysis is intended to be of interest for stakeholders and policymakers to design and implement forms of collaboration with similar regions across southern Europe that may have the same needs in terms of the enhancement of their innovation structures. Then we identified the two indicators with the greatest impact for each of the four categories of indicators, applying the attribute-related sensitivity analysis, according to poset theory. As two categories already consisted of two indicators, we conducted the analysis just for the most numerous categories, namely ‘framework conditions’ and ‘innovation activities’. Then we performed the poset-based analysis on the final dataset, which consists of the scores on the most impacting indicators of the nine clusters, and after the data analysis we obtained the ranking. The cluster that was found to be at the top level consists of the five Greek regions of Attiki, Dytiki Ellada, Kentriki Makedonia, Kriti, and Thessalia, and two Portuguese regions of Centro, and Lisboa. The indicators for which the top-level cluster is the best are *R&D expenditure in public sector as percentage of GDP* and *SMEs introducing marketing or organisational innovations as percentage of SMEs*. The positive results for Greece for public R&D investments also find support in the literature; in fact, Zoumpekas et al. (2021) analysing the data of the European Innovation Scoreboard from 2010 to 2017, demonstrated that Greece significantly enhanced its score for this indicator compared to the EU average. The middle-bottom and bottom levels of our analysis mostly consist of the remaining Greek regions and most regions of central and southern Spain. Regarding the Italian regions, they are placed at the middle-top level (most of northern and central Italy) and the middle level (mainly the southern regions).

The second phase of the analysis revealed more detailed information about the top-performing regions for each country. The leader in Portugal is Lisbon; the Italian top level consists of Friuli-Venezia Giulia, Provincia Autonoma di Trento, and Toscana; the leader in Greece is Kriti; four regions are top performing in Spain: Cataluña, Comunidad Valenciana, Madrid, and País Vasco. The results of the second phase also showed the ranking obtained in the first phase of the analysis to be consistent.

We then identified the most impacting indicators resulting from the combination of the five attribute-related sensitivity analyses that we conducted in our work (once in the first phase, and four times in the second phase). We found that in the category ‘framework conditions’, the indicator *top-10% most cited publications worldwide as percentage of total scientific publications of the country* was most impacting four times, followed by the indicator *percentage of population aged 30-34 with tertiary education* (three times); whereas for the category ‘innovation activities’, the indicators that resulted most impacting were *SMEs introducing marketing or organisational innovations as percentage of SMEs*, and *European design applications per billion GDP in PPS* (four times each). The attribute-related sensitivity analysis also enables us to ascertain that there are indicators that are most impacting only for some countries, as in the case of the indicator *SMEs innovating in-house as percentage of SMEs*, most impacting for Portugal, or the indicator *international scientific co-publications per million population*, most impacting for Italy. Regions can concentrate their efforts to improve the results for these indicators to improve their position in the ranking.

Finally, we compared the results of the first phase of the poset-based analysis with the performance categories presented in the RIS 2019, and we found that 29 out of 60 regions analysed are at the same performance level in both analyses, and that 26 other regions change position by just one level up or down. Three regions (Dytiki Ellada, Isles Baleares, and Thessalia), according to our analysis, improved their performance significantly; conversely, the outcomes for Aragón and Cantabria in the poset-based analysis compared well to the RIS 2019. The difference is explained by high (in the case of the first three regions) or low (in the case of the two Spanish regions) performances for the most impacting indicators compared to the average of all indicators, which is adopted by the RIS to construct the ranking.

For future research, it would be interesting to apply the poset-based analysis also to the data of the Regional Innovation Scoreboard relating to the previous years, to find the trend of innovation in southern Europe over the last decade. Furthermore, the same study could also be conducted on the national data of the European Innovation Scoreboard.

The coming years will be crucial for the whole of Europe, especially for the southern European regions, thanks to the opportunities of the Recovery Plan and of the Next Generation EU programmes. For this reason, understanding the patterns of innovation could help to relaunch the economies and to foster resilience.

Appendix A. Regional Innovation Scoreboard 2019

This is the ninth edition of the Regional Innovation Scoreboard (RIS) (the first one was published in 2009) and provides a comparative assessment of the performance of regional innovation systems across 238 regions of 23 EU Member States, together with Norway, Serbia, and Switzerland. The RIS is associated with the European Innovation Scoreboard (EIS), which assesses the performance of national innovation systems. The RIS assigns the European regions to four innovation performance groups: *innovation leaders* (i.e., regions with a relative performance greater than 120% of the EU average), *strong innovators* (i.e., regions with a relative performance between 90% and 120% of the EU average), *moderate innovators* (i.e., regions with a relative performance between 50% and 90% of the EU average), and *modest innovators* (i.e., regions with a relative performance below 50% of the EU average).

The RIS aims to measure innovation performance using the same 27 indicators adopted by the EIS, though regional data are not available for many indicators. As a result, the RIS assesses regions considering 17 indicators grouped into four different categories: 1. framework conditions (*1a. percentage of population aged 30-34 with tertiary education*, *1b. lifelong learning* – the share of population aged 25-64 enrolled in education or training aimed at improving knowledge, skills and competences, *1c. international scientific co-publications per million population*, *1d. top-10% most cited publications worldwide as percentage of total scientific publications of the country*); 2. investments (*2a. R&D expenditures in public sector as percentage of GDP*, *2b. R&D expenditures in business sector as percentage of GDP*, *2c. SMEs non-R&D innovation expenditures as percentage of GDP*); 3. innovation activities (*3a. SMEs introducing product or process innovations as percentage of SMEs*, *3b. SMEs introducing marketing or organisational innovations as percentage of SMEs*, *3c. SMEs innovating in-house as percentage of SMEs*, *3d. innovative SMEs collaborating with others as percentage of SMEs*, *3e. public-private co-publications per million population*, *3f. PCT patent applications per billion GDP in PPS*, *3g. trademark applications per billion GDP in PPS*, *3h. European design applications per billion GDP in PPS*); 4. impacts (*4a. employment in medium-high and high-tech manufacturing and knowledge-intensive services as percentage of total employment*, *4b. SMEs sales of new-to-market and new-to-firm innovations as percentage of total turnover*).

Most of the data relating to the listed indicators are obtained from the Community Innovation Survey data with the help of National Statistical Offices, and from Eurostat. Other sources include the Centre for Science and Technology Studies of Leiden University, Science Metrics, and the OECD's REGPAT database. The missing data are imputed adopting a range of techniques based on the availability of regional or national data referring to the previous years of observation. In relation to certain indicators of some regions for which data are difficult to find, there are still missing data even after imputation; however, data availability increases to approximately 99% (Hollanders et al., 2019b). The data are then normalised following the min-max procedure: the minimum normalised score is equal to 0 and the maximum normalised score is equal to 1. The final regional score is obtained as the unweighted average of the 17 indicators multiplied with a country correction factor (Hollanders et al., 2019b).

An interesting aspect of the RIS 2019 is the average score of the indicators per regional performance group. Considering the EU average equal to 100, the report of the RIS 2019 shows that 15 out of 17 indicators have the best score in the innovation leaders group and the worst score in the modest innovators group. Just two indicators follow a different pattern. The first one is the indicator related to *innovative SMEs collaborating with others*, which has a slightly higher score in the strong innovators group than in the innovation leaders group (126 vs 118). However, the difference is small, and in the moderate and modest innovator groups the score is much lower compared to the innovation leaders group. The second one is the indicator related to *non-R&D innovation expenditures*, which is the only one in which moderate innovator regions have a performance that is higher than 100% of the EU average, and the outcome of the innovation leaders group is similar to that of the modest innovators. Hence, it seems that in this context this indicator does not respect the outcomes of the innovation performance groups. As explained in the methodological report of the RIS 2019, the strong performance of both moderate and modest innovators on *non-R&D innovation expenditures* reflects the fact that in less innovative regions, it is more cost-effective for enterprises to innovate by purchasing advanced machinery and equipment, and knowledge developed elsewhere, than to invest in their own R&D activities as they are more expensive and at higher risk of failing to result in a useful product or process innovation (Hollanders et al., 2019b). The issues about the non-R&D innovation expenditures indicator have already been discussed in the literature, for instance, in Blažek & Kadlec (2019) as well as in Spescha & Woerter (2019). For this reason, we excluded this indicator from our analysis.

Table A.1 shows the distribution of the 60 southern European regions over the performance groups according to the RIS 2019 report.

Table A.1 – The performance groups of the 60 regions of southern Europe, according to the RIS 2019 report

Performance Group (RIS 2019)	N° of regions
Strong Innovators	5
Moderate Innovators +	12
Moderate Innovators	24
Moderate Innovators -	13
Modest Innovators	6

Considering the regions represented in Table A.1, five regions are categorised as strong innovators (Lisboa, Norte, and Centro in Portugal, Kriti in Greece, and Friuli-Venezia Giulia in Italy), six regions are categorised as modest innovators (Canarias, Castilla-la Mancha, Extremadura, Ceuta and Melilla in Spain, and Notio Aigaio – or Southern Aegean - in Greece), and all the other 49 regions are categorised as moderate innovators. Hence, the Moderate group is by far the largest one considering Greece, Italy, Portugal and Spain; however, three sub-categories are proposed for each performance group in the RIS 2019 report. By virtue of this, 12 regions are classified as "moderate innovators +", 24 regions are classified as simply "moderate innovators", and the remaining 13 regions are classified as "moderate innovators –". The same results can be shown by highlighting the number of regions of each country belonging to the different performance groups, as reported in Table A.2.

Table A.2 – Performance group memberships of the southern European regions country by country, according to the RIS 2019 report

Country	**Greece**	**Italy**	**Portugal**	**Spain**
Strong	1	1	3	-
Moderate +	2	8	-	2
Moderate	6	7	4	7
Moderate -	3	5	-	5
Modest	1	-	-	5

According to the results of the RIS 2019, the 60 regions of Greece, Italy, Portugal and Spain are grouped in five different performance groups. On the one hand, Portugal and Italy have no regions classified as modest innovators; on the other hand, Spain has no regions considered as strong innovators; finally, most of the regions (24 out of 60) are in the moderate innovators group, in the middle of this ranking.

Appendix B – Poset-based approach

To illustrate the characteristics of the theory of partially ordered sets, we provide a simple example as a guide for the analysis performed in this paper.

Consider four given objects a, b, c and d, three indicators q_1, q_2, and q_3, and the average of the indicators μ, as described in Table B.1. We will call the set of objects X, and the set of indicators A. In the table below, we consider a scenario with three numerical indicators in which the higher the score, the better the outcome. However, in poset theory, indicators are just features and they could also be linguistic descriptions (i.e. high, medium, low, etc.) or ordinal indicators.

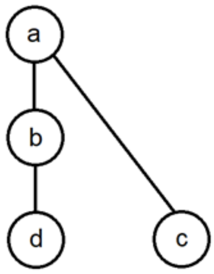
Table B.1 – Example: objects, indicators and the average of the indicators

Objects	q_1	q_2	q_3	μ
a	6	3	3	4
b	3	2	2	2.3
c	5	1	2	2.7
d	2	2	1	1.7

If we simply calculate the average of all indicators to determine the ranking, we will easily find that object a leads the ranking with a score of 4, followed by object c (2.7), and finally objects b and d (respectively with a score of 2.3 and 1.7). However, using the average may result in misleading conclusions. In the poset analysis this is avoided, since it is crucial to compare all the objects based on all attributes. Therefore, we could say that object a (6,3,3) is better than object b (3,2,2), object c (5,1,2) and object d (2,2,1) since it shows a higher score on all attributes. We could also say that object b is better than object d because even if the two objects tie on q_2 (2 for both objects b and d), object b has a higher score on both q_1 and q_3 compared to object d (3 for object b and 2 for object d on q_1 and 2 for object b and 1 for object d on q_3). What is not possible to compare is object c with objects b and d: c shows a higher score on q_1 compared to both objects b and d ($5 > 3$ and $5 > 2$), but a lower score on q_2 ($1 < 2$); hence, object c is comparable with object a only, and incomparable with objects b and d.

Looking at the dataset, we could then establish the relations between the comparable objects: a > b > d, as well as a > c. At the same time, we know that c || b and c || d (where || is the sign to represent incomparability). The result can be also represented through a Hasse diagram, as in Figure B.1.

Figure B.1 – Example: Hasse diagram



Now it is possible to identify the downset and the upset of any of the objects. The downset of an object *x* consists of those objects *y* such that $y \leq x$; its cardinality is denoted as D(x). If $y < x$ for one or more indicators and $y > x$, then *x* and *y* are incomparable; the number of objects that are incomparable with an object *x* is denoted as I(x). We obtain Table B.2.

Table B.2 - Example: downsets and incomparabilities of the objects, in numbers

Objects	D(x)	I(x)
a	4	0
b	2	1
c	1	2
d	1	1

In Table B.2 it is possible to see, for instance, that the downset of object b consists of two objects (objects b and d).

We are now able to rank the objects of the poset. The method adopted is the Local Partial Order Model (LPOM), where the "final score" of an object is a function of D(x) and I(x). The formula to compute the "final score" $\delta(x)$ of any object is as follows: (Brüggemann & Patil, 2011)

$$\delta(x) = \mathrm{D}(x)\,[(n+1)/\,(n+1-\mathrm{I}(x))] \qquad \text{(B.1)}$$

where *x* is the object of interest and *n* indicates the total number of objects, which, in this case, is $n = 4$. For instance, the score of object a, applying the formula, is: 4 * (4 + 1) / (4 + 1 – 0) = 4 * 5 / 5 = 4. After computing the score for all the objects, we obtain the following ranking: a, b, c, d; which is different from the ranking obtained by simply calculating the average of the indicators, which in this case yields a, c, b, d. Hence, the Hasse diagram highlights which objects are without doubt better (or worse) than the others. With the LPOM it is possible to rank all the objects, even if some of them are incomparable.

Finally, in the poset-based analysis, it is possible to reduce the number of attributes through the "attribute-related sensitivity" analysis. The aim is to examine how an attribute influences the position of the objects in the Hasse diagram by removing a column from the data matrix (Brüggemann & Patil, 2011). The goal, now, is to find the pair of attributes (out of three) that makes it possible to reproduce the original Hasse diagram of Figure B.1.

We first have to identify the downset of each object considering the whole data matrix (X, A). Then we compare these identified downsets with the ones of all objects (X) considering the same data matrix with the exclusion of one attribute at a time. To find, for instance, the impact of q_1, we have to look at the columns (X, A) and (X, A\{q_1}) of Table B.3: for each object, we identify what are the downsets considering the two different data matrices. We can see in Table B.3 that the downset of object b in (X, A) consists of two objects (b and d), but it consists of three objects in (X, A\{q_1}) (objects b, c and d). The total difference in cardinality between the two data matrices (counting the number of objects that form the downsets) is 1, as indicated in the last row of Table B.3. We then repeat the same exercise excluding indicators q_2 and q_3.

Table B.3 - Example: attribute-related sensitivity analysis. Downsets of the objects in X for different subsets of attributes

Objects	(X, A)	(X, A\{q_1})	(X, A\{q_2})	(X, A\{q_3})
a	{a, b, c, d}	{a, b, c, d}	{a, b, c, d}	{a, b, c, d}
b	{b, d}	{b, c, d}	{b, d}	{b, d}
c	{c}	{c}	{b, c, d}	{c}
d	{d}	{d}	{d}	{d}
Total difference in cardinality		1	2	0

As shown in Table B.3, q_3 has no impact on the results, while excluding attribute q_2 results in two differences. In fact, without q_2, object c is higher than both objects b and d, which is not the case in the original data matrix (in Table B.3 the differences are marked in red). Finally, it is possible to conclude that the pair of attributes that best represents the original Hasse diagram is formed by q_1 and q_2, therefore if we intend to simplify the data matrix, we can consider just the first two indicators.

Appendix C

This appendix lists the 60 regions included in the study. All the regions are listed in Table C.1 following the alphabetic order of their NUTS 2 code (first column). The second column presents the name of the region, whereas the columns from the third to the sixth show respectively the number of clusters (according to the first phase of the analysis), the level of the first phase of the analysis, the level of the second phase of the analysis (country-level focuses), and the group in the Regional Innovation Scoreboard 2019.

Table C.1 – Results of the 60 southern European regions, according to the poset-based analysis and the Regional Innovation Scoreboard 2019 report

Region	Cluster	Level (1st phase)	Level (2nd phase)	Group RIS 2019
Abruzzo	9	Middle	Middle-bottom	Moderate
Alentejo	3	Middle-bottom	Bottom	Moderate
Algarve	3	Middle-bottom	Middle	Moderate
Anatoliki Makedonia, Thraki	1	Middle-bottom	Bottom	Moderate-
Andalucía	4	Bottom	Middle-bottom	Moderate-
Aragón	4	Bottom	Middle	Moderate
Attiki	2	Top	Middle-top	Moderate+
Basilicata	9	Middle	Middle-bottom	Moderate
Calabria	9	Middle	Bottom	Moderate-
Campania	9	Middle	Middle-bottom	Moderate
Canarias	4	Bottom	Middle-bottom	Modest+
Cantabria	4	Bottom	Middle	Moderate
Castilla y León	4	Bottom	Middle-bottom	Moderate-
Castilla-la Mancha	4	Bottom	Middle-bottom	Modest+
Cataluña	5	Middle-top	Top	Moderate+
Centro	2	Top	Middle-top	Strong-
Ciudad Autónoma de Ceuta	7	Bottom	Bottom	Modest-
Ciudad Autónoma de Melilla	4	Bottom	Middle-bottom	Modest
Comunidad de Madrid	5	Middle-top	Top	Moderate
Comunidad Foral de Navarra	5	Middle-top	Top	Moderate

Comunidad Valenciana	6	Middle-top	Middle-top	Moderate
Dytiki Ellada	2	Top	Middle-top	Moderate
Dytiki Makedonia	1	Middle-bottom	Middle	Moderate
Emilia-Romagna	8	Middle-top	Middle	Moderate+
Extremadura	4	Bottom	Middle-bottom	Modest+
Friuli-Venezia Giulia	8	Middle-top	Top	Strong-
Galicia	4	Bottom	Middle	Moderate-
Ionia Nisia	1	Middle-bottom	Bottom	Moderate
Ipeiros	3	Middle-bottom	Middle	Moderate
Isles Baleares	6	Middle-top	Middle-top	Moderate-
Kentriki Makedonia	2	Top	Middle-top	Moderate+
Kriti	2	Top	Top	Strong-
La Rioja	6	Middle-top	Middle-top	Moderate
Lazio	9	Middle	Middle-top	Moderate
Liguria	9	Middle	Middle-top	Moderate
Lisboa	2	Top	Top	Strong-
Lombardia	8	Middle-top	Middle	Moderate+
Marche	8	Middle-top	Middle	Moderate+
Molise	9	Middle	Middle-bottom	Moderate-
Norte	8	Middle-top	Middle	Strong-
Notio Aigaio	3	Middle-bottom	Middle-bottom	Modest
País Vasco	5	Middle-top	Top	Moderate+
Peloponnisis	1	Middle-bottom	Bottom	Moderate-
Piemonte	8	Middle-top	Middle	Moderate+
Principado de Asturias	4	Bottom	Middle	Moderate-
Provincia Autonoma Bolzano	9	Middle	Middle	Moderate
Provincia Autonoma Trento	8	Middle-top	Top	Moderate+
Puglia	9	Middle	Middle-bottom	Moderate
Região Autónoma dos Açores	3	Middle-bottom	Middle-bottom	Moderate
Região Autónoma da Madeira	9	Middle	Middle	Moderate

Región de Murcia	6	Middle-top	Middle-top	Moderate
Sardegna	9	Middle	Bottom	Moderate-
Sicilia	9	Middle	Bottom	Moderate-
Sterea Ellada	1	Middle-bottom	Bottom	Moderate
Thessalia	2	Top	Middle-top	Moderate
Toscana	8	Middle-top	Top	Moderate+
Umbria	8	Middle-top	Middle	Moderate+
Valle d'Aosta	9	Middle	Middle	Moderate-
Veneto	8	Middle-top	Middle	Moderate+
Voreio Aigaio	3	Middle-bottom	Middle	Moderate-

References

Alaimo, L.S., Arcagni, A., Fattore, M. & Maggino, F. (2020). Synthesis of Multi-indicator System Over Time: A Poset-based Approach. *Social Indicators Research*. doi:10.1007/s11205-020-02398-5.

Arbolino, R., Boffardi, R. & De Simone, L. (2019). Which are the Factors Influencing Innovation Performances? Evidence from Italian Cohesion Policy. *Social Indicators Research 146*, 221-247. doi:10.1007/s11205-018-1904-5.

Asheim, B.T., Lawton Smith, H. & Oughton, C. (2011). Regional Innovation Systems: Theory, Empirics and Policy. *Regional Studies 45 (7)*, 875-891. doi:10.1080/00343404.2011.596701.

Asheim, B.T., Isaksen, A. & Trippl, M. (2019). Advanced Introduction to Regional Innovation Systems. Edward Elgar ISBN: 978 1 78536 197 5.

Badinger, H. & Reuter, W.H. (2015). Measurement of fiscal rules: Introducing the application of partially ordered set (POSET) theory. *Journal of Macroeconomics 43*, 108-123. doi:10.1016/j.jmacro.2014.09.005.

Balland, P.A. & Boschma, R. (2021). Complementary interregional linkages and Smart Specialisation: an empirical study on European regions. *Regional Studies (online)*. doi:10.1080/00343404.2020.1861240. https://www.tandfonline.com/doi/full/10.1080/00343404.2020.1861240, last accessed 25 January 2021.

Bekiros, S., Hammoudeh, S., Jammazi, R. & Nguyen, D.K. (2018). Sovereign bond market dependencies and crisis transmission around the eurozone debt crisis: a dynamic copula approach. *Applied Economics 50 (47)*, 5031-5049. doi:10.1080/00036846.2018.1470313.

Blažec, J. & Kadlec, V. (2019). Knowledge bases, R&D structure and socio-economic and innovation performance of European regions. Innovation: *The European Journal of Social Science Research 32 (1)*, 26-47. doi:10.1080/13511610.2018.1491000.

Brüggemann, R. & Patil, G.P. (2011). Ranking and Prioritization for Multi-Indicator Systems. Introduction to Partial Order Applications. Springer-Verlag New York. ISBN: 978-1-4419-8476-0. doi:10.1007/978-1-4419-8477-7.

Capello, R. & Lenzi, C. (2019). Regional innovation evolution and economic performance. *Regional Studies 53 (9)*, 1240-1251. doi:10.1080/00343404.2018.1502421.

Carayannis, E.G., Goletsis, Y. & Grigoroudis, E. (2018). Composite innovation metrics: MCDA and the Quadruple Innovation Helix framework. *Technological Forecasting and Social Change 131*, 4-17. doi:10.1016/j.techfore.2017.03.008.

Carlsen, L. (2018). Happiness as a sustainability factor. The world happiness index: a posetic-based data analysis. *Sustainability Science 13*, 549-571. doi:10.1007/s11625-017-0482-9.

Coenen, L., Asheim, B., Bugge, M.M. & Herstad, S.J. (2017). Advancing regional innovation systems: What does evolutionary economic geography bring to the policy table? *Environment and Planning C: Politics and Space 35 (4)*, 600-620. doi:10.1177/0263774X16646583.

De Loof, K., De Baets, B., De Meyer, H. & Brüggemann, R. (2008). A hitchhiker's guide to poset ranking. *Combinatorial chemistry & high throughput screening 11 (9)*, 734-744. doi:10.2174/138620708786306032.

Doloreux, D. & Parto, S. (2004). Regional Innovation Systems: A Critical Review. *MERIT Working Paper*. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.477.377&rep=rep1&type=pdf, last accessed 25th January 2021.

Doloreux, D. & Porto Gomez, I. (2017). A review of (almost) 20 years of regional innovation systems research. *European Planning Studies 25(3)*, 371-387. doi:10.1080/09654313.2016.1244516.

Donatiello, D. & Ramella, F. (2017). The Innovation Paradox in Southern Europe. Unexpected Performance During the Economic Crisis. *South European Society and Policies 22 (2*), 157-177. doi:10.1080/13608746.2017.1327339.

Fattore, M. & Arcagni, A. (2014). PARSEC: An R Package for Poset-Based Evaluation of Multidimensional Poverty. In: *Multi-indicator Systems and Modelling in Partial Order 317-330. Springer, New York, NY*. ISBN: 978-1-4614-8222-2. doi:10.1007/978-1-4614-8223-9_15.

Fattore, M., Maggino, F. & Colombo, E. (2012). From Composite Indicators to Partial Orders: Evaluating Socio-Economic Phenomena Through Ordinal Data. In: *Quality of life in Italy. Social Indicators Research Series 41-68. Springer, Dordrecht*. ISBN: 978-94-007-3897-3. doi:10.1007/978-94-007-3898-0_4.

Franco, S., Gianelle, C., Kleibrink, A. & Murciego, A. (2020). Learning from similar regions: how to benchmark innovation systems beyond rankings. *Chapter 7* 162-194 in Capello et al., Quantitative Methods for Place-Based Innovation Policy. doi:10.4337/9781789905519. ISBN: 9781789905502, pp 256.

Gabriel, K. (2019). Accelerating innovation in a changing world. *International Journal of Research, Innovation and Commercialisation 2 (2)*, 105. doi:10.1504/IJRIC.2019.104014.

Garcia-Bernabeu, A., Cabello, J.M. & Ruiz, F. (2020). A Multi-Criteria Reference Point Based Approach for Assessing Regional Innovation Performance in Spain. *Mathematics 8 (5)*, 797. doi:10.3390/math8050797.

Garofalo, A., Castellano, R., Punzo, G. & Musella, G. (2018). Skills and labour incomes: how unequal is Italy as part of the Southern European countries? *Quality & Quantity 52*, 1471-1500. doi:10.1007/s11135-017-0531-6.

Grillitsch, M. & Asheim, B. (2018). Place-based innovation policy for industrial diversification in regions. *European Planning Studies 26 (8)*, 1638-1662. doi:10.1080/09654313.2018.1484892.

Hauser, C., Siller, M., Schatzer, T., Walde, J. & Tappainer, G. (2018). Measuring regional innovation: A critical inspection of the ability of single indicators to shape technological change. *Technological Forecasting and Social Change 129*, 43-55. doi:10.1016/j.techfore.2017.10.019.

Hollanders, H., Es-Sadki, N. & Merkelbach, I. (2019a). Regional Innovation Scoreboard 2019. *Publications Office of the European Union*. ISBN:978-92-76-08723-6. doi:10.2873/89165.

Hollanders, H., Es-Sadki, N. & Merkelbach, I. (2019b). Regional Innovation Scoreboard 2019. Methodology Report. *European Union*. Available at https://ec.europa.eu/docsroom/documents/37783, last accessed 4 February 2021.

Jadhav, A., Pramod, D. & Ramanathan, K. (2019). Comparison of Performance of Data Imputation Methods for Numeric Dataset. *Applied Artificial Intelligence 33 (10)*, 913-933. doi:10.1080/08839514.2019.1637138.

Lau, A.K.W. & Lo, W. (2015). Regional innovation system, absorptive capacity and innovation performance: An empirical study. *Technological Forecasting and Social Change 92*, 99-114. doi:10.1016/j.techfore.2014.11.005.

Mazzucato, M. (2018). Mission-Oriented Research & Innovation in the European Union. A problem-solving approach to fuel innovation-led growth. *Publications Office of the European Union*. ISBN: 978-92-79-79918-1. doi:10.2777/36546.

Moreira, A., Léon, M., Coda Moscarola, F. & Roumpakis, A. (2020). In the eye of the storm…again! Social policy responses to COVID19 in Southern Europe. *Social & Policy Administration (online)*. doi:10.1111/spol.12681. https://onlinelibrary.wiley.com/doi/full/10.1111/spol.12681, last accessed: 25 January 2021.

Morrison, A., Doussineau, M. (2019). Regional innovation governance and place-based policies: design, implementation and implications. *Regional Studies, Regional Science 6 (1)*, 101-116. doi:10.1080/21681376.2019.1578257.

Navarro, M., Gibaja, J., Bilbao-Osorio, B. & Aguado, R. (2009). Patterns of innovation in EU-25 regions: a typology and policy recommendations. *Environment and Planning C: Government and Policy 27 (5)*, 815-840. doi:10.1068/c0884r.

Nunes, S., Carvalho, L. & Costa, T. (2013). Cooperation for innovation: evidence from southern European countries. *International Journal of Innovation and Regional Development 5 (2), 226-241*. doi:10.1504/IJIRD.2013.055250.

Perez, S.A. & Matsaganis, M. (2018). The Political Economy of Austerity in Southern Europe. *New Political Economy 23 (2)*, 192-207. doi:10.1080/13563467.2017.1370445.

Ponsiglione, C., Quinto, I. & Zollo, G. (2018). Regional Innovation Systems as Complex Adaptive Systems: The Case of Lagging European Regions. *Sustainability 10 (8), 2862*. doi:10.3390/su10082862.

Rodil-Marzábal, O. & Vence-Deza, X. (2020). Regional Innovation Systems and regional disparities in the Euro area: insights for regional innovation policy. *Chapter 7*, 139-161, in González-López, M. & Asheim, B.T., Regions and Innovation policies in Europe, learning form the margins. New horizons in regional sciences series. doi:10.4337/9781789904161.00012. ISBN:9781789904154.

Rodríguez-Pose, A., Ketterer, T. (2020). Institutional change and the development of lagging regions in Europe. *Regional Studies 54 (7)*, 974-986. doi:10.1080/00343404.2019.1608356.

Russell, M.G. & Smorodinskaya, N.V. (2018). Leveraging complexity for ecosystemic innovation. *Technological Forecasting and Social Change 136*, 114-131. doi:10.1016/j.techfore.2017.11.024.

Spescha, A. & Woerter, M. (2019). Innovation and firm growth over the business cycle. *Industry and Innovation 26 (3)*, 321-347. doi:10.1080/13662716.2018.1431523.

Tsakovski, S., Astel, A. & Simeonov, V. (2010). Assessment of the water quality of a river catchment by chemometric expertise. *Journal of Chemometrics 24 (11-12)*, 694,702. doi:10.1002/cem.1333.

Zabala-Iturriagagoitia, J.M., Voigt, P., Gutiérrez-Gracia, A. & Jiménez-Sáez, F. (2007). Regional Innovation Systems: How to Assess Performance. *Regional Studies 41 (5)*, 661-672. doi:10.1080/00343400601120270.

Zoumpekas, T., Vavalis, M., & Houstis, E. (2021). Analysis of innovation with data science: The case of Greece. International Journal of Data Science and Big Data Analytics 1 (1), 20-42. doi:10.51483/IJDSBDA.1.1.2021.20-42.