Received: 2 November 2018 | Revised: 15 October 2019 | Accepted: 24 October 2019

DOI: 10.1111/pirs.12492

FULL ARTICLE

Relatedness in the implementation of Smart Specialisation Strategy: a first empirical assessment

Diego D'Adda[1] | **Donato Iacobucci**[1] | **Roberto Palloni**[2]

[1] Centre for Innovation and Entrepreneurship & Dipartimento di Ingegneria dell'Informazione, Università Politecnica delle Marche, Ancona, Italy

[2] t33, via Calatafimi 1, 60121 Ancona, Italy

Correspondence
Donato Iacobucci, Centre for Innovation and Entrepreneurship & Dipartimento di Ingegneria dell'Informazione, Università Politecnica delle Marche, Via Brecce Bianche, 12-60131. Ancona, Italy.
Email: d.iacobucci@univpm.it

Abstract

Smart Specialisation Strategy required regional authorities to identify technological domains where to concentrate investment in R&D and innovation. Their choices should originate from an analysis of regional strengths, aiming to identify those domains with the greatest potential for innovation and diversification. This paper aim to assess the degree of relatedness between chosen technological domains. We use patent data to categorize technological domains and the revealed associations methodology to measure their degree of relatedness. We find that the choices made by Italian regions show a significant higher relatedness as compared to a random choice, though with notable exceptions.

KEYWORDS

regional development, regional innovation policy, smart specialisation strategy, technological relatedness

JEL CLASSIFICATION

O38; R11; R58

1 | INTRODUCTION

The EU Cohesion Policy is one of the western world's largest, if not the largest, regional development policy operating under broadly one overall legal and institutional framework. Following the so-called Lisbon strategy, the main aim of the 2014–2020 program is to foster the innovative performance of EU regions and promoting a better link between the production of new knowledge, as resulting from R&D investment, and its application to new products and services.[1] To achieve this aim, the EU has developed the concept of the Smart Specialisation Strategy (S3) (Foray, David, & Hall, 2009; McCann & Ortega-Argilés, 2013; McCann & Ortega-Argilés, 2015). The design of S3 is an ex-ante conditionality for EU regions for the allocation of structural and investment funds in the 2014–2020 programming period.[2] The design of S3 required national and regional authorities to identify technological domains where to concentrate private and public investment in R&D and innovation (Foray, 2015; Foray et al., 2009). The concept of S3 is based on two fundamental ideas: a) a region should not spread its investments in too many different fields and focus them on few technological domains (specialisation); b) these domains have to be chosen in order to enhance or complement the research and productive assets the region is already endowed with (smart) (Capello, 2013; Foray et al., 2012; Foray & Goenega, 2013). The design of the S3 introduced two main novelties. The first is the emphasis on the 'entrepreneurial discovery'; this implies a bottom up approach in the identification of the specialisation fields, based on a systematic consultation of regional stakeholders. The second important novelty is that regions are required to identify technological domains rather than industry sectors. Targeting technological domains instead of specific industries is expected to enhance product innovation and diversification by creating new technologies and new productions (Asheim & Grillitsch, 2015; Foray, David, & Hall, 2011).

The choice of the specialisation domains should originate from an analysis of the region's actual strengths and weaknesses, aiming to identify the domains with the greatest potential for innovation and diversification. With this idea in mind, the EU guidelines for S3 design explicitly mention the concept of relatedness as one of the main criteria to take into account in choosing the specialisation domains (Foray et al., 2012). In doing so, the S3 guide recognizes the importance of related variety by considering the 'cross-fertilization' of ideas between different technological domains as a key factor to promote innovation (especially product innovation) and diversification (Asheim, Boschma, & Cooke, 2011; Frenken, Van Oort, & Verburg, 2007; Grillitsch, Asheim, & Trippl, 2018).

Despite the importance attributed to related variety in S3 design, the analysis of S3 documents approved by EU regions reveals that only a few regions explicitly considered the relatedness between technological domains as a criteria for their specialisation choices (Capello & Kroll, 2016; Iacobucci, 2014; Iacobucci & Guzzini, 2016). This can be a potential weakness in the implementation of the S3 since technological relatedness at regional level is considered a key factor for innovation and diversification (Asheim et al., 2011; Boschma & Frenken, 2011a; Lambooy & Boschma, 2001; Neffke, Henning, & Boschma, 2011).

In general, the operationalization of S3 has been rather limited because of the lack of a consolidated set of analytical tools to be applied by regional authorities in the design and implementation of the strategy. Another reason is that specialisation domains are indicated by regions using the natural language, which reduces comparability and hinder the possibility to perform quantitative analysis.

To overcome these problems, in order to carry out an empirical analysis we first have associated the specialisation domains declared in the S3 documents with the corresponding International Patent Classification (IPC) codes. Then, we used the revealed associations methodology to measure the degree of relatedness of technological domains and asses to what extent the choices made by regions have fulfilled the S3 requirement, specifically about the level of relatedness between technological domains.

More precisely, our analysis conducted on 19 Italian regions followed these steps:

1. assignment of the domains chosen by each region to an international and standardized technology classification (International Patent Classification),
2. calculation of a product space based on European patent data and of proximity measures between technological domains "à la Hidalgo" (see below for a detailed explanation),

[1]For an overview of the policy and the budget involved see: https://cohesiondata.ec.europa.eu/

[2]The resources allocated within the ESIF (European Structural and Investment Funds) are about 638 Billion Euros. A significant share is explicitly devoted to sustaining research and innovation.

3. calculation of relatedness measures based on proximities for
    a. the technological domains chosen by regions in S3,
    b. the technological domains in which regions are actually specialised,
4. comparison the relatedness measures for chosen and actual domains.

The paper makes contributions at the methodological and empirical level. At the methodological level we suggest an indicator to measure the degree of relatedness between the technological domains chosen by regions. On the empirical side, we perform a first evaluation about whether regional choices fulfilled the criteria suggested for the design of S3. Our methodology provides a measure of relatedness that can help regions to refine their specialisation choices. Moreover, this measure can also be used in the ex-post evaluation of the policy in order to capture the regional heterogeneity in the design of S3 policy. In fact, any assessment exercise of the effectiveness of this policy must consider regional differences in its design and implementation. Italy is particularly suited for this empirical analysis given the large number of regions and their heterogeneity in terms of size and level of development.

The paper is organized as follows. Section 2 discusses the importance of relatedness in the implementation of S3. Section 3 presents the data and the methodology at the basis of our indicators. Section 4 discusses the empirical results obtained by applying the analysis to Italian regions. Section 5 draws the conclusions and suggests some policy implications.

2 | S3 AND THE CONCEPT OF RELATEDNESS: THEORETICAL ARGUMENTS AND PRACTICAL ISSUES

The adoption of the Smart Specialisation Strategy (S3) was an ex-ante conditionality for EU regions for the allocation of the structural and investment funds in the 2014–2020 programming period. S3 is one of the major novelties in EU innovation policy. It is based on the idea that regions should concentrate resources in a limited number of technological domains in which they are more likely to achieve a significant impact in terms of innovation and technological diversification (Foray, 2015; Foray et al., 2009).

The S3 guide (Foray et al., 2012) emphasizes that regions should consider the potential relations between the specialisation domains in designing their strategy: “S3 aims at developing world class excellence clusters and providing arenas for related variety and cross sectoral links which drive specialised technological diversification” (Foray et al., 2012, p. 6).

The emphasis on related variety in S3 policy is not surprising since technological relatedness has become a central concept in the literature about innovation and regional development. The presence of related variety within the same region is expected to provide two main benefits: promoting innovation through the cross fertilization of knowledge between different sectors (Boschma & Frenken, 2011a; Frenken et al., 2007); favoring the process of diversification into new sectors (Boschma & Frenken, 2011b; Neffke et al., 2011).

The concept of related variety has been emphasized by several authors discussing the S3 rationale. McCann and Ortega-Argilés (2015) sustain that regions should specialise in different “*knowledge-related sectors*”. This is relevant because: “*domains that are highly connected with other domains will offer greater possibilities for learning than less connected domains*” (McCann and Ortega-Argilés, 2015, p. 3). The authors underline the importance of related variety as a way to achieve technological diversification, especially for those peripheral regions that show an excessive reliance on few technological domains. Similarly, Boschma and Gianelle (2014) agree that related variety between technological domains is not only beneficial to foster innovation performance and growth but also for the diversification of the regional industrial base.

Despite the importance acknowledged to the relatedness between specialisation domains, applying the concept of ‘related variety’ in the S3 design is not easy as it raises several questions both at the theoretical and practical level (Boschma, 2014). At the theoretical level, the related variety approach could clash with the ‘critical mass principle’, which is the main underlying rationale of a specialisation strategy. This problem is particularly relevant for small regions which may face difficulties in investing in too many technological domains at the same time. In fact, the ‘related variety approach’ is based hinges on the ‘Jacobian’ agglomeration advantages, which are mostly observed in rich (and large) urban contexts (Duranton & Puga, 2001; Jacobs, 1969).

Besides the theoretical questions, there are several problems arising at an empirical level when trying to define and measure knowledge relations between sectors. Boschma and Gianelle (2014) suggest some methods that regions could use to measure the degree of relatedness between sectors: inter-industry relatedness, based on the classification codes of economic activities; co-occurrence of products; the presence of input–output linkages; the intensity of labor reallocation between industries. The latter method, based on the revealed skill relatedness (RSR) proposed by Neffke and Henning (2013), is suggested also in the S3 guide (Foray et al., 2012, p. 37).

Up to now, most of the theoretical and empirical works on related variety have taken into consideration products and industries rather than technological domains (Boschma & Iammarino, 2009; Boschma, Minondo, & Navarro, 2010; Hartog, Boschma, & Sotarauta, 2012). The specialisation domains identified within the S3 are not expected to be too wide (as an entire industry) or to narrow (an individual firm). They should refer to an intermediate level which involve groups of firms, preferably belonging to different industry sectors, with the aim of exploring new domains of technological and market opportunities which are relevant for the region (Asheim & Grillitsch, 2015; Foray, 2015; Foray & Goenega, 2013). In practical terms, this means identifying technological domains at the intersection of different industries. Moreover, technological domains are chosen not only considering their relative strength within the region but also thinking about their potential application to new activities. Indeed, the entrepreneurial discovery process is also aimed at exploiting the presence of knowledge spillovers which are at the basis of the creation of new cluster of activities (Foray & Goenega, 2013). This aspect is emphasized in recent contributions. Balland and Rigby (2017) suggest that in implementing S3 regions should not only support the development of related areas of potential specialisation but also the development of “related technologies that are more complex that what they already produce” (Balland & Rigby, 2017, p. 19).

Despite the fact that the importance of targeting related domains in the implementation of S3 is widely acknowledged, only a minority of regions addressed the question of relatedness in choosing the specialisation domains and none of them attempted to use specific measures of relatedness (Capello & Kroll, 2016; Iacobucci, 2014; Iacobucci & Guzzini, 2016). There are two main reasons for this fact. The first is the absence of a common methodology to assess the degree of relatedness between technological domains. The second is the absence of suitable data to apply the available methodologies. For example, the methodology suggested in the S3 guide, based on revealed skill relatedness (RSR), would require the availability of individual data on labor mobility between firms, which are available only for a limited number of EU regions. Another obstacle is that regions expressed specialisation domains using the natural language (e.g.: “biotech”, “health and wellness”, “mechatronics”, etc.).[3] In general, the use of the natural language reduces comparability and limits the possibility to perform quantitative analysis.

We overcome the latter problem by defining the specialisation domains chosen by regions using the International Patent Classification (IPC). Using IPC codes to analyze relatedness has also the advantage to allow the identification of technological domains rather than industry sectors, as required by S3 logic. The homogenization of specialisation domains, i.e. adopting a common classification system, made it possible to perform an empirical analysis aimed at measuring the degree of relatedness between the specialisation domains chosen by regions in their S3. This is then compared to the actual degree of relatedness between technological domains already present in the region. Moreover, it also allowed us to assess to what extent regions were able to choose technological domains that show a higher degree of relatedness with respect to the level of relatedness between the technological domains in which the region is actually specialised.

[3]This engenders some drawbacks. First, there is little control about the content that is included under a specific label. Second, different labels could be used to refer to the same technological domains.

TABLE 1 An example of association of IPC codes to detailed technologies: the Aerospace domain in Campania

Region	S3 domain	Detailed description of technologies	IPC Class
Campania	Aerospace	Technologies and production processes for the development of metallic structures	C21
		Development of advanced ceramic materials for the foundry of precision	B28
		Development and tuning of systems of high-pressure injection	F02
		Guidance systems for advanced autonomous navigation and control	G05
		…	…
	Total aerospace	35 different technologies	13 distinct IPC classes

Source: S3 documents approved by Campania regional authorities further elaborated by us.

3 | DATA AND METHODOLOGY

In this paper, we measure the degree of relatedness between technological domains chosen by regional authorities and we compare it to the relatedness between technological domains in which regions are actually specialised. The empirical analysis is based on two main sources of data. In order to track the chosen specialisation domains we used the S3 documents officially approved by Italian regions. Instead, to compute technological relatedness we used patent data retrieved from the OECD RegPat database.

3.1 | S3 data

From the S3 documents officially approved by 19 Italian regions we have retrieved the description of the technological domains chosen by the regional authorities at the highest level of detail. Since every region has listed them using natural language and without referring to a common classification system, we first had to homogenize the taxonomy in order to have fully comparable information. To this aim, we performed a systematic association between the description of technological domains provided in the S3 documents and the corresponding International Patent Classification (IPC) classes.[4] This association was carried out in a semi-automated way by using the publicly available service IPCCAT (Categorization Assistant in the International Patent Classification).[5] Then, this automatic mapping has been manually revised by experts.

For example, the Campania region has indicated 6 "upper-level" technological domains among which Aerospace. Under the latter domain, the S3 document listed 35 specific technological sub-domains (at the highest level of detail). An IPC code was associated to each of them, resulting in 13 unique different IPC classes. Therefore, we obtained a detailed map of the chosen technological domains and of the corresponding IPC classes.[6] This specific example is illustrated in Table 1.

It is worthwhile noting that also when regions chose traditional sectors as target of their S3, (e.g. agri-food, tourism or culture), the detailed description of technological (sub)domains often refers to technologies applied to those sectors; mostly ICT and other key enabling technologies. For example, the Lazio region has indicated the "Artistic and cultural industries" as priority target. After providing a list of the specific industries – film, music, design and architecture, fashion, arts, etc. – the document states that the specialisation domain refers to the application of digital technologies to those industries. As a result, most of the specific technologies indicated within this target sector

[4]www.wipo.int/portal/en/. Please note the an IPC class is identified by a code composed by one letter and two digits, .e.g. A21 identifies "baking; equipment for making or processing doughs; doughs for baking".

[5]www.wipo.int/ipccat/

[6]Abruzzo is the only Italian region not included in the analysis because the S3 document initially approved indicated the general domains of specialization (namely, agrifood, life science, ICT/aerospace, fashion/design) but did not provide a detailed description of the technologies included in these domains. This made it impossible to determine the IPC classes associated to the specialization domains.

TABLE 2 An example: the technologies listed within the specialization domain "Cultural heritage and technology for culture" by the Lazio region

S3 domain	Detailed description of technologies
Cultural heritage and technology for culture	remote sensing, photogrammetric surveys, excavation technologies and archaeological research technologies for the documentation and cataloging of the territory micro and macro climate monitoring materials, diagnostics and advanced instrumentation for restoration and conservation; digitization and cataloging of cultural heritage semantic web for cultural heritage open source and open data web infrastructures data mining e-story-telling lighting technology advanced filmography (e.g. special effects, 3D, project mapping) gamification mobility apps virtual museums and virtual tours governance models, museum and cultural heritage management and related business models

Source: S3 document approved by Lazio region.

belong to the ICT. The same logic applies also to the domain labelled "Cultural heritage": the Lazio region has indicated a series of advanced technologies for the diagnosis, restoration, conservation and fruition of cultural heritage (see Table 2). For this reason, we were able to associate the technologies as indicated by regions to the corresponding IPC classes.[7]

Overall, the specialisation domains listed in the S3 documents were mapped into 64 different IPC codes, with G06 (Data processing systems or methods) being the class with the highest frequency across regions. Out of the 64 IPC codes, 20 codes (31%) have a frequency equal to 1 (i.e. they are indicated by one region only), suggesting a moderate level of diversification of the Italian regions in choosing their specialisation domains. The choice of IPC codes to characterize specialisation domains is crucial to detect similarities or differences in regional choices. Indeed, some regions have included different technologies under the same label or used different labels to refer to the same technological domain.

As a result, each Italian region k is characterized by a set of IPC codes corresponding to the technological domains chosen within its S3 strategy. In other words, if an IPC code i is in the set then the region k has indicated the corresponding technological domains in its S3 documents.

3.2 | Patent data

We retrieved patent data using the OECD RegPat database (February 2016 version) which provides information about the IPC codes a patent belongs to and the address of its applicant(s) and inventor(s). We looked at the PCT patent applications from 2002–2012 with at least one of the inventors localized in Europe. This choice was made in order to assign to each European NUTS2 region the corresponding number of patents. These data allow us to assess the dynamics of innovation and technological specialisation within the European regions as measured by patents. While the shortcomings of patent data as a measure of innovation are well known, they are widely and increasingly used at the regional level to measure knowledge flow and knowledge specialisation (see e.g. Kogler, Essletzbichler, & Rigby, 2017).

[7]Out of about 1200 specific technologies we were not able to classify about 20% or the items. In most cases, this was not because they referred to non-technological areas but because the definition was too broad/imprecise, such as for example the label "smart materials".

We computed the fractional count of PCT applications for each IPC class (if a patent was classified in more than one IPC class, its fractional count is considered for every IPC class it belongs to) in each year and in every European NUTS2 region.

3.3 | Measures and indicators

Our relatedness indicators are based on the comparison between the specialisation domains chosen by regions and the proximity measures computed using patents data. In order to build the proposed indexes, we followed the steps listed below:

1. we compute the specialisation for 271 European regions in every IPC Class (122 IPC classes), hence obtaining a 271 × 122 specialisation matrix,
2. we compute the proximity measure à la Hidalgo, that is a measure of the probability that a region is specialised in a specific pair of IPC classes (hence obtaining a 122x122 proximity matrix),
3. taking into consideration the technological domains as *chosen* by regional authorities and represented with IPC classes, we compute two indexes based on the average proximity between every possible combination of those IPC classes (namely the *Average Relatedness Index* and the *Relatedness Share Index*),
4. taking into consideration the technological domains in which regions are *actually specialised* (again represented with IPC classes), we compute two indexes based on the average proximity between every possible combination of those IPC classes (namely the *Average Relatedness Index Actual* and the *Relatedness Share Index Actual*).

In the following paragraphs we will describe in detail all these steps.

3.4 | Proximity measures

We propose different measures of relatedness based on different ways of computing the 'proximity' between technological domains. In particular, we make use of two "geographical/macro-level" measures based on Hidalgo, Klinger, Barabasi, and Hausmann (2007), one based on relative specialisation and another one on absolute specialisation.[8]

1. Proximity à la Hidalgo based on relative specialisation measures

In order to measure the actual relative technological specialisation at the NUTS2 level we used the Balassa Index, also known as Revealed Comparative Advantage (RCA) index. We defined the RCA as follows

$$\mathrm{RCA}_{k,i} = \frac{\frac{X_{ki}}{\sum_{k=1}^{K} X_{ki}}}{\frac{\sum_{i=1}^{I} X_{ki}}{\sum_{k=1,i=1}^{K,I} X_{ki}}}.$$

X_{ki} is the sum of the fractional count of PCT patents in the period 2002–2012 in region k and belonging to IPC class i.[9] Therefore, $RCA_{k,i}$ is the ratio between the patent share of region k in IPC class i and the patent share of IPC class i in the world. Since the Balassa index tends to have an asymmetric and skewed distribution, we computed a symmetric version of it by applying the following transformation (Dalum, Laursen, & Villumsen, 1998):

[8]Please note that other measures of proximity have been proposed by the literature, e.g. a patent / micro-level measure based on the co-occurrence of IPC classes within the same patent document (also known as co-classification codes, see e.g. Breschi, Lissoni, & Malerba, 2003). We replicated our analysis using this alternative measure and results are available from the authors upon request.

[9]Please note that in order to compute proximity measures we compute the specialisation in each IPC class (122) for 271 NUTS2 regions in Europe.

$$\mathrm{RCA}_{\mathrm{k,i}}^{\mathrm{norm}} = \frac{\mathrm{RCA}_{\mathrm{k,i}} - 1}{\mathrm{RCA}_{\mathrm{k,i}} + 1}.$$

We dichotomized the $RCA_{k,i}^{norm}$ using the threshold 0, since values below 0 point at a negative relative specialisation in a certain technological domain as identified by the IPC class of a region, while values above 0 indicate a positive relative specialisation.

Following Hidalgo et al. (2007), we compute a measure of proximity between any pair of IPC classes *i* and *j* by taking the minimum of the pairwise conditional probability (using the European regions as observations)[10]:

$$proximity_{i,j} = min\left(P\left(RCA_i > 0 | RCA_j > 0\right), P\left(RCA_j > 0 | RCA_i > 0\right)\right),$$

where

$$\mathrm{P}\left(\mathrm{RCA_i} > 0 | \mathrm{RCA_j} > 0\right) = \frac{\mathrm{P}\left(\mathrm{RCA_i} > 0 \wedge \mathrm{RCA_j} > 0\right)}{\mathrm{P}\left(\mathrm{RCA_j} > 0\right)}.$$

The proximity matrix is a $N \times N$ symmetric matrix containing the revealed technological proximity between any two IPC classes *i* and *j*. Each cell (*i,j*) represents the probability that a (random) region that is relatively specialised in *i*(*j*) is specialised in *j*(*i*) as well. Therefore, this matrix is composed of the revealed proximities between technological domains. The matrix is symmetric and diagonal elements are arbitrary assigned a value equal to 0. The underlying idea is that when there is a frequent association between specialisations in different technological classes, it should be optimal to combine (at the geographical level) those specialisations in producing new technological knowledge (e.g. patents). We report in Appendix A some statistics to give a sense of the proximity matrix and its structure.

2. Proximity à la Hidalgo based on absolute value measures

Besides the measure of relative specialisation, we also take into account the absolute "importance" of a region in a certain IPC class by considering the fractional count of patents in the period 2002–2012. The variable $NPat_{k,i}$ is equal to the fractional count of PCT applications of the region *k* in the IPC class *i* over the period 2002–2012. This measure identifies the number of patents with at least one of its inventors located in a specific European NUTS2 region.[11] Then, we dichotomize this variable using the first quartile of nonzero values across regions as threshold, namely *t*. The use of this absolute measure is more in line with the idea of related variety since the exchange of knowledge between sectors depends on the absolute stock of resources invested in those sectors rather than their relative importance.

Again, following Hidalgo et al. (2007), we compute a measure of proximity between any pair of IPC classes *i* and *j* taking the minimum of the pairwise conditional probability (using the European regions as observations):

$$proximity_{i,j} = min\left(P\left(NPat_i > t | NPat_j > t\right), P\left(NPat_j > t | NPat_i > t\right)\right),$$

where

$$\mathrm{P}\left(\mathrm{NPat_i} > \mathrm{t} | \mathrm{NPat_j} > \mathrm{t}\right) = \frac{\mathrm{P}\left(\mathrm{NPat_i} > \mathrm{t} \wedge \mathrm{NPat_j} > \mathrm{t}\right)}{\mathrm{P}\left(\mathrm{NPat_j} > \mathrm{t}\right)}.$$

The proximity matrix is a $N \times N$ symmetric matrix containing the revealed technological proximity between any two IPC classes *i* and *j*. Each cell (*i,j*) represents the probability that a (random) region that is specialised (in absolute terms) in *i*(*j*) is specialised in *j*(*i*) as well.

[10]Please note that this measure is agnostic as to the reasons why two specializations (either product or technological) should co-occur at the local level. In fact, it just looks at the probability of co-occurrence without explaining or suggesting the underlying reason driving the co-localization pattern.

[11]Please note that we compute the absolute specialisation in each IPC class (122) for 271 NUTS2 regions in Europe.

3.5 | Relatedness indicators

In this paper we aim to assess to which extent Italian regions have chosen S3 technological domains with a high degree of relatedness. In order to do so, we compute several indexes aimed at capturing the degree of relatedness between the specialisation domains chosen by regions. In particular, for each region *k* we compute the following indexes:[12]

$$\mathrm{ARI_k} = \frac{\sum_{\mathrm{i} \in \mathrm{C_k}} \sum_{\mathrm{j} \in \mathrm{C_k},\, \mathrm{j} \neq \mathrm{i}} proximity_{\mathrm{i,\,j}}}{\sum_{\mathrm{i} \in \mathrm{C_k}} \sum_{\mathrm{j} \in \mathrm{C_k},\, \mathrm{j} \neq \mathrm{i}} 1}$$

and

$$\mathrm{RSI_k} = \frac{\sum_{\mathrm{i} \in \mathrm{C_k}} \sum_{\mathrm{j} \in \mathrm{C_k},\, \mathrm{j} \neq \mathrm{i}} \mathrm{I}\left(proximity_{\mathrm{i,\,j}} > median\right)}{\sum_{\mathrm{i} \in \mathrm{C_k}} \sum_{\mathrm{j} \in \mathrm{C_k},\, \mathrm{j} \neq \mathrm{i}} 1}.$$

$\mathrm{proximity_{i,j}}$ is the proximity between the technological domains (i.e. IPC codes) *i* and *j*.

Please note that they can be computed using alternatively the two proximity measures we illustrated in the previous subsection, namely: 1) the proximity à la Hidalgo based on relative specialisation measures, 2) the proximity à la Hidalgo based on absolute specialisation measures. The first index, the Average Relatedness Index (ARI), is basically an average of the proximities between any pairs of IPC codes chosen by each region *k* (i.e. belonging to $\mathrm{C_k}$, the group of technological domains chosen by region *k*). The second index, namely the Relatedness Share Index (RSI), represents the percentage of pairs of IPC codes chosen by each region *k* that have proximity above a threshold. In particular, $\mathrm{I}(\mathrm{proximity_{i,j}} > median)$ is an indicator function equal to 1 when the proximity is above the median of proximities distribution.

Last but not least, we also compare the chosen technological specialisation at the regional level with the actual observed technological specialisation. In particular, for each region *k* we compute the following indexes:[13]

$$\mathrm{ARI}_k^{actual} = \frac{\sum_{\mathrm{i} \in \mathrm{S_k}} \sum_{\mathrm{j} \in \mathrm{S_k},\, \mathrm{j} \neq \mathrm{i}} proximity_{\mathrm{i,\,j}}}{\sum_{\mathrm{i} \in \mathrm{S_k}} \sum_{\mathrm{j} \in \mathrm{S_k},\, \mathrm{j} \neq \mathrm{i}} 1}$$

and

$$\mathrm{RSI}_k^{actual} = \frac{\sum_{\mathrm{i} \in \mathrm{S_k}} \sum_{\mathrm{j} \in \mathrm{S_k},\, \mathrm{j} \neq \mathrm{i}} \mathrm{I}\left(proximity_{\mathrm{i,\,j}} > median\right)}{\sum_{\mathrm{i} \in \mathrm{S_k}} \sum_{\mathrm{j} \in \mathrm{S_k},\, \mathrm{j} \neq \mathrm{i}} 1}.$$

These two indexes are analogous to the previously mentioned ones. Please note that ARI_k^{actual} is an average of the proximities between any pairs of IPC codes in which the region *k* is actually specialised. Specifically, $\mathrm{S_k}$ is the group of technological domains in which the region *k* is specialised, in relative terms (as measured by $RCA_{k,i}^{norm} > 0$), or in absolute term (as measured by $NPat_i > t$). Again, $proximity_{\mathrm{i,j}}$ can be measured as: 1) the proximity à la Hidalgo based on relative specialisation measures, 2) the proximity à la Hidalgo based on absolute specialisation measures. Please note that in case a region chose the same technological domains in which it is actually specialised, then its ARI (RSI) would be equal to its $\mathrm{ARI^{actual}}$ ($\mathrm{RSI^{actual}}$). Moreover, this choice would be consistent with the principle of embeddedness as well (D'Adda, Guzzini, Iacobucci, & Palloni, 2018; McCann & Ortega-Argilés, 2015), i.e. targeting technological

[12]By way of example, let's hypothesize that a region has chosen in its S3 strategy as priority technological domains (IPC codes) A21, A22 and A01. In addition, let's say that $\mathrm{proximity_{A21,A22}} = \mathrm{proximity_{A22,A21}} = 0.2$ and $\mathrm{proximity_{A01,A21}} = \mathrm{proximity_{A21,A01}} = 0.6$ and $\mathrm{proximity_{A01,A22}} = \mathrm{proximity_{A22,A01}} = 0.7$. Therefore, $\mathrm{ARI_k} = \frac{1.5}{3} = 0.5$.

[13]By way of example, let's hypothesize that a region is actually specialised in technological domains (IPC codes) A21, A22 and A01. In addition, let's say that $\mathrm{proximity_{A21,A22}} = \mathrm{proximity_{A22,A21}} = 0.2$ and $\mathrm{proximity_{A01,A21}} = \mathrm{proximity_{A21,A01}} = 0.6$ and $\mathrm{proximity_{A01,A22}} = \mathrm{proximity_{A22,A01}} = 0.7$. Therefore, $\mathrm{ARI_k} = \frac{1.5}{3} = 0.5$.

domains according to the actual regional capabilities. Therefore, in case a region has carefully and strategically selected as target few technological domains looking at i) the actual strengths and ii) the level of relatedness between these domains, we expect its ARI (RSI) to be higher than its ARI^{actual} (RSI^{actual}).

4 | EMPIRICAL RESULTS

Before analyzing the relatedness indicators at regional level, Table 3 reports the correlations between the indexes, and between the indexes and the number of technological domains chosen by regions (*n_chosen*), the number of actual specialisations in relative (*n_spec*) and absolute (*n_absspec*) terms, the GDP (*gdp*), the GDP per capita (*gdp_percapita*), the population (*pop*) and the innovativeness of the region (*innoindex*).[14] We added these additional variables in order to check whether our variables and indicators are somehow correlated with the economic characteristics of a region.

Reasonably, bigger and more innovative regions should have chosen a higher number of technological domains[15] within their S3, though the correlation is not significant. All the indexes are slightly negatively correlated with the number of chosen domains, although without significance. This is not surprising given the construction of our indexes as an average of the all pairs of chosen IPC. In principle, such an index can be maximized by taking the (unique) pair of IPC with the highest proximity value. Adding any other IPC (less close to the initial pair by definition) would lower this index. We may have a similar effect also when a region has two clusters with a high level of proximity within each cluster but a relative low proximity between the two clusters. In this case the index would be lower than in the case of having one single cluster of related technological domains. In this sense, caeteris paribus the indexes are likely to be lower for bigger regions with respect to smaller ones, but this fact is absolutely connected to the very definition of relatedness.[16] Thus, it is not surprising that the GDP is positively related to the number of actual specialisations while being negatively correlated with the ARI computed using the proximity based on the absolute specialisation.

The correlation table suggests also a very high correlation between ARI and RSI, both using relative and absolute specialisation indexes to measure proximity between technological domains. The ARI is probably more precise while the RSI is easier to interpret; the high correlation between the two highlights that they are almost interchangeable for the empirical analysis.

Finally, we found a rather high correlation between the indexes computed using proximity measures based on relative and on absolute specialisation. This is particularly true for the two ARI indexes, which show a correlation close to 0.7; this means that the indexes show a low level of sensitivity to different formulations.

Using the proximity measures based on the relative specialisation, Table 3 reports the number of technological domains chosen by Italian regions, the number of actual specialisations, the ARI and the RSI indexes computed as described in the previous section, based on the choices made by regions and based on the actual specialisations. Table 4 is analogous, but it uses the proximity measures based on the absolute specialisation instead.

The number of technological domains chosen within the S3 is quite different across regions, with an average number of 13 but ranging from 5 to 33 IPC classes. This is not surprising given the large differences in the size of Italian regions, form less than 200 thousand people in Valle d'Aosta to about 10 million in Lombardy.

The mean value of the ARI_r (Average Relatedness Index) for Italian regions (0.32) is above the average proximity between technological domains (0.26) observed in EU regions (i.e. the expected proximity if two domains were chosen at random).

[14]The data source for the GDP, the GDP per capita, the population is Eurostat (2016) while for the innovativeness of the region is the Regional Innovation Scoreboard (2017).

[15]Please note that we are measuring the declared intentions, not the actual diversification patters of regions. A higher number of chosen technological domains points to the intention (by regional authorities) to achieve a higher level of diversification while the average relatedness index (ARI) takes this intention as given in order to understand the degree of relatedness of regional choices.

[16]However, please note that the ARI and the RSI are not driven by the number of chosen domains "by construction". In fact, randomly simulating the indexes as a function of the number of chosen domains and looking at the distribution of simulated values, the mean and median are stable.

TABLE 3 Correlation matrix

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	***n_chosen***		0.350	−0.437	−0.289	−0.122	−0.096	0.293	−0.340	−0.136	−0.014	−0.112	0.319	0.001	0.408	0.154
2	***n_spec***	0.350		−0.355	−0.244	0.005	0.175	0.840***	−0.233	−0.307	−0.530*	−0.498*	0.739***	0.367	0.682***	0.589**
3	$\boldsymbol{ARI_r}$	−0.437	−0.355		0.951***	0.139	−0.049	−0.227	0.694***	0.402	−0.105	−0.007	−0.331	−0.079	−0.347	−0.048
4	$\boldsymbol{RSI_r}$	−0.289	−0.244	0.951***		0.090	−0.083	−0.158	0.571**	0.309	−0.052	0.006	−0.287	−0.172	−0.269	−0.067
5	$\boldsymbol{ARI^{actual}{}_r}$	−0.122	0.005	0.139	0.090		0.920***	−0.150	0.186	0.175	0.321	0.329	−0.194	0.049	−0.189	0.108
6	$\boldsymbol{RSI^{actual}{}_r}$	−0.096	0.175	−0.049	−0.083	0.920***		−0.011	0.025	0.019	0.236	0.184	−0.117	0.184	−0.149	0.235
7	***n_absspec***	0.293	0.840***	−0.227	−0.158	−0.150	−0.011		−0.110	−0.252	−0.746***	−0.652**	0.741***	0.405	0.695***	0.708***
8	$\boldsymbol{ARI_a}$	−0.340	−0.233	0.694***	0.571**	0.186	0.025	−0.110		0.839***	−0.232	−0.124	−0.076	−0.029	−0.119	−0.041
9	$\boldsymbol{RSI_a}$	−0.136	−0.307	0.402	0.309	0.175	0.019	−0.252	0.839***		0.048	0.099	−0.069	−0.025	−0.124	−0.097
10	$\boldsymbol{ARI^{actual}{}_a}$	−0.014	−0.530*	−0.105	−0.052	0.321	0.236	−0.746***	−0.232	0.048		0.892***	−0.643**	−0.427	−0.543*	−0.511*
11	$\boldsymbol{RSI^{actual}{}_a}$	−0.112	−0.498*	−0.007	0.006	0.329	0.184	−0.652**	−0.124	0.099	0.892***		−0.559*	−0.409	−0.495*	−0.374
12	***gdp***	0.319	0.739***	−0.331	−0.287	−0.194	−0.117	0.741***	−0.076	−0.069	−0.643**	−0.559*		0.302	0.941***	0.401
13	***gdp_percapita***	0.001	0.367	−0.079	−0.172	0.049	0.184	0.405	−0.029	−0.025	−0.427	−0.409	0.302		0.053	0.686***
14	***pop***	0.408	0.682***	−0.347	−0.269	−0.189	−0.149	0.695***	−0.119	−0.124	−0.543*	−0.495*	0.941***	0.053		0.195
15	***innoindex***	0.154	0.589**	−0.048	−0.067	0.108	0.235	0.708***	−0.041	−0.097	−0.511*	−0.374	0.401	0.686***	0.195	

TABLE 4 Relatedness indicators using the proximity measure *à la* Hidalgo based on relative specialization measures

			Indicators based on S3 regional choices				Indicators based on actual regional technological specialization	
Region	Number of IPC codes chosen by the region in the S3 (n_chosen)	Number of IPC codes in which the region show a positive relative specialization (n_spec)	Average relatedness index (ARI_r)	Fisher testp–value (with respect to a random choice)	Fisher testp–value (with respect to actual specialization domains)	Relatedness share index (RSI_r)	Average relatedness index (ARI^{actual}_r)	Relatedness share index (RSI^{actual}_r)
Basilicata	13	29	0.224	0.952	1.000	32%	0.292	61%
Calabria	17	39	0.397	0.000***	0.005***	90%	0.335	72%
Campania	33	47	0.276	0.354	1.000	49%	0.304	61%
Emilia–Romagna	15	52	0.277	0.391	1.000	46%	0.345	84%
Friuli Venezia Giulia	16	43	0.335	0.005***	0.514	62%	0.336	79%
Lazio	17	45	0.284	0.270	0.931	46%	0.316	62%
Liguria	16	50	0.339	0.004***	0.018**	71%	0.296	61%
Lombardia	19	70	0.264	0.627	0.999	43%	0.310	67%
Marche	10	52	0.293	0.242	0.755	51%	0.317	66%
Molise	5	25	0.415	0.003***	0.088*	80%	0.336	68%
Piemonte	15	59	0.301	0.113	0.706	57%	0.314	68%
P.A. Bolzano	9	44	0.322	0.070*	0.792	53%	0.347	82%
P.A. Trento	22	48	0.284	0.247	1.000	54%	0.354	82%
Puglia	11	46	0.275	0.439	0.996	45%	0.335	76%
Sardegna	8	44	0.298	0.216	0.875	54%	0.334	78%
Sicilia	10	40	0.328	0.041**	0.627	64%	0.340	74%
Toscana	6	48	0.385	0.009***	0.021**	80%	0.297	60%
Umbria	10	46	0.345	0.012**	0.395	60%	0.337	76%
Valle d'Aosta	9	25	0.363	0.006***	0.031**	64%	0.304	62%
Veneto	13	56	0.346	0.005***	0.251	65%	0.331	76%
mean	13.70	45.40	0.32			58.33%	0.32	70.72%
std.dev.	6.35	10.78	0.05			0.14	0.02	0.08

TABLE 5 Relatedness indicators using the proximity measure *à la* Hidalgo based on absolute specialization measures

			Indicators based on S3 regional choices				Indicators based on actual regional technological specialization	
Region	Number of IPC codes chosen by the region in the S3 (n_chosen)	Number of IPC codes in which the region show a positive absolute specialization (n_absspec)	Average relatedness index (ARI_a)	Fisher testp–value (with respect to a random choice)	Fisher testp–value (with respect to actual specialization domains)	Relatedness share index (RSI_a)	Average relatedness index ($ARI^{actual}{}_a$)	Relatedness share index ($RSI^{actual}{}_a$)
Basilicata	13	2	0.690	0.160	–	74%	0.823	100%
Calabria	17	21	0.814	0.000***	0.021**	88%	0.776	77%
Campania	33	76	0.754	0.000***	0.064*	74%	0.717	69%
Emilia–Romagna	15	113	0.801	0.000***	0.000***	83%	0.667	57%
Friuli Venezia Giulia	16	90	0.749	0.012***	0.204	66%	0.718	71%
Lazio	17	106	0.803	0.000***	0.000***	97%	0.686	60%
Liguria	16	82	0.783	0.001***	0.014**	78%	0.710	66%
Lombardia	19	119	0.740	0.014**	0.024**	69%	0.646	53%
Marche	10	80	0.713	0.128	0.501	51%	0.707	72%
Molise	5	2	0.865	0.001***	–	100%	0.740	100%
Piemonte	15	111	0.755	0.010**	0.043**	72%	0.673	58%
P.A. Bolzano	9	39	0.792	0.009***	0.269	83%	0.772	85%
P.A. Trento	22	44	0.688	0.096*	1.000	66%	0.795	92%
Puglia	11	62	0.703	0.144	0.786	45%	0.740	74%
Sardegna	8	33	0.793	0.011**	0.140	86%	0.740	69%
Sicilia	10	42	0.766	0.019**	0.635	80%	0.779	86%
Toscana	6	111	0.783	0.037**	0.074*	67%	0.661	58%
Umbria	10	46	0.852	0.000***	0.002***	100%	0.730	73%
Valle d'Aosta	9	6	0.786	0.012**	–	78%	0.676	47%
Veneto	13	112	0.785	0.002***	0.004***	85%	0.672	59%
mean	13.70	64.85	0.77			77.1%	0.72	71.3%
std.dev.	6.35	40.39	0.05			0.14	0.05	0.15

To provide a measure of the ability of regions to choose related sectors, we performed an adaptation of the Fisher's exact test[17] in order to understand the "distance" between the ARI resulting from the regional choices and one resulting from a random choice of technological domains. Specifically, the reported p-value is the probability that a random extraction of IPC classes would result in an ARI equal or greater than the observed one; thus, giving an indication of the "goodness" (or distance) of the choice of a region with respect to a random selection. We perform a similar test simulating the distribution of the ARI that would result by randomly choosing the IPC classes within the set of IPC classes in which the region is actually specialised. In other words, we estimate the probability that a random choice within the technological domains in which the region is specialised would result in a lower relatedness with respect to the one based on the observed choices.[18] The first test highlights the presence of large differences in the behavior of regions. For some of them, like Calabria, Friuli Venezia Giulia, Liguria, Molise, Tuscany and Veneto, the ARI_r is significantly higher compared to a random choice. In others, the test reveals that the value of the relatedness index is not statistically different from the value that would have resulted if the regional authorities had selected the technological domains present in their region at random. The value of the test is logically related to the value of RSI, i.e. the percentage of the chosen domains that show a high level of technological relatedness. The second test is more restrictive, leading to significant differences in fewer cases. For example, Umbria's ARI is significantly different from an ARI based on a completely random choice (i.e. randomly picking *n_chosen*, i.e. 10, IPC sectors out of 122) but it is not different once we take into consideration a random choice within the technological domains in which it is actually specialised (i.e. randomly picking *n_chosen*, i.e. 10, IPC sectors out of *n_spec*, i.e. 46). This may be a result of Umbria having an actual ARI higher than average and choosing between the technological domains in which the region is actually specialised but without taking into consideration the relatedness itself. For example, with respect to Umbria, Toscana choices seem to achieve a higher degree of relatedness even considering its actual specialisations and this fact points to a higher adherence to S3 guidelines/requirements in terms of related diversification.

The situation changes significantly, both in terms of the ARI and RSI, when we consider the indices based on absolute rather than relative specialisation (see Table 5). As mentioned in the previous section, the absolute specialisation can be considered more close to the concept of related variety at local level: i.e. the actual possibility of exchanging knowledge and resources between different technological domains. Also, in this case (as previously observed for ARI_r) the mean value of the ARI_a in Italian regions (0.72) is above the average proximity between technological domains observed in EU regions (0.68). Even more important, for most regions, the choices in terms of technological domains show a significantly higher degree of relatedness with respect to a random choice.

Considering the ability of regions to select related technological domains, we have to acknowledge that there is a trade-off between the maximization of relatedness (i.e. choosing domains that show a high degree of relatedness) and the maximization of the coherence between the chosen domains and those in which the region is actually specialised (see D'Adda et al., 2018). Both these criteria were requirements in S3 guidelines. For this reason, we computed the relatedness indicators in the hypotheses that each region had chosen the IPC codes of the technological domains in which it is actually specialised. The actual Average Relatedness Index and the actual Relatedness Share Index are computed as described in the previous section, using the proximity measures based on the relative and the absolute specialisation, but considering the actual specialisation of regions rather than the choices made by regional authorities for S3. These indexes give an indication about the degree of relatedness of the technological domains in which a region is actually specialised (relatively or absolutely). Looking at the correlation between these indexes (see Table 2), we see again a very high correlation between ARI and RSI.

[17]In particular, we performed 10,000 random drawings of *n* IPC classes, with *n* equal to the number of IPC codes chosen by the region in the S3. For each drawing, we computed the ARI. We then compare the distribution of the 10,000 simulated ARIs to the value of ARI we actually observe (the one coming from the choices actually made by regions). The reported p-value is the probability that a random extraction of IPC would result in an ARI equal or greater than the observed one. Please note that the higher *n*, the lower the variance of the distribution of the simulated ARI, the narrower the distribution and the closer to the mean value.

[18]This Fisher test compares the degree of relatedness of actual choices with the degree that would have been obtained picking *n_chosen* domains at random among the n IPC classes in which the region is actually specialised. Please note that a region may have picked technological domains outside of those in which it's actually specialised, but this is a matter of coherence (i.e. using S3 words, embeddedness).

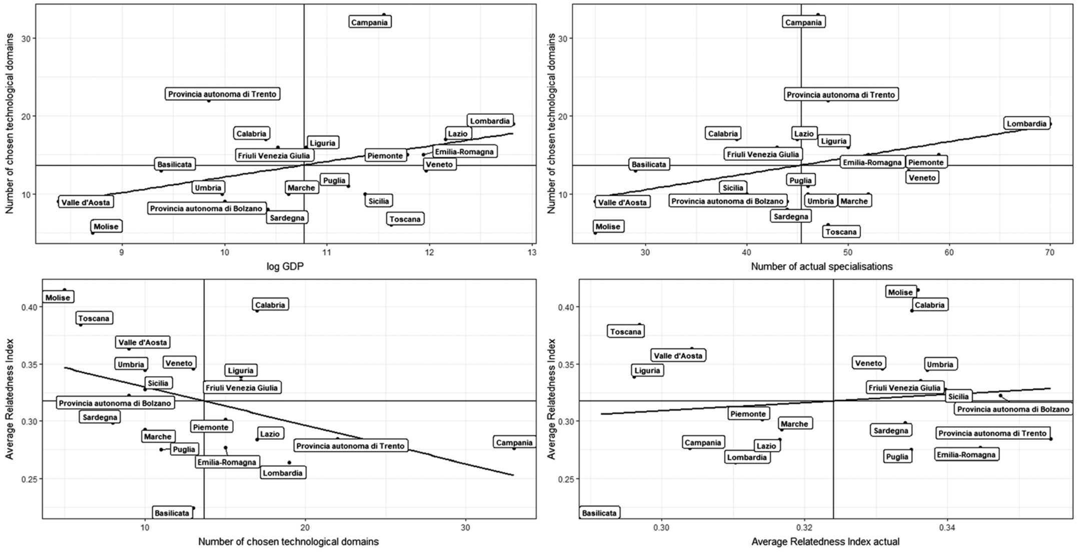

FIGURE 1 – Representing regional position in terms of relatedness indicators (based on relative specialisation)

An interesting fact is that the indexes measuring actual relatedness are not correlated with the ones measuring relatedness of the technological domains chosen by Italian regions. In principle, a region should choose technological domains i) in which it is specialised (i.e. it has competences and know-how) and ii) that are closely related to each other (i.e. with high proximity). Therefore, if regions had chosen with these criteria in mind, we would expect the indexes based on actual specialisation to be lower than the indexes based on the chosen domains. The difference between the two set of indexes is close to zero on average. This means that regions have probably privileged the technological domains in which they already show a high degree of specialisation rather than trying to maximize the relatedness between the chosen domains.

In Figure 1 we present some graphs aimed to show the relative position of Italian regions along different dimensions. The first two graphs highlight that the number of chosen domains seem somehow positively related with the GDP and the number of actual specialisation in a region. A notable exception is Campania region, having it chosen a seemingly disproportionate number of technological domains in which (not) to specialise. The bottom left graph shows a slightly negative relation (though not significant) between the ARI and the number of chosen technological domains. Basilicata and Campania regions seem to differ sharply from the average, with the former being characterized by a lower than average number of chosen domains and very low ARI, and the latter having a very high ARI notwithstanding an above -average number of chosen domains. Lastly, the bottom right graph highlights the scarce correlation between the ARI based on chosen technological domains and the ARI based on actual specialisations, with the regions spread quite uniformly in the four quadrants.

5 | CONCLUSIONS AND FURTHER DEVELOPMENT

The Smart Specialisation Strategy applied by the EU for the allocation of structural funds for the programming period 2014–2020 required regions to choose a set of technological domains in which to concentrate R&D investments and innovation policies. Among other important novelties, S3 embraced the idea recently developed within the economic geography literature that relatedness between technological domains may promote innovation and facilitate diversification. As a result, the EU guidelines for S3 required regions to choose the specialisation domains considering not only their actual specialisation but also the degree of relatedness between these domains.

Despite the emphasis on this latter concept, only few regions mentioned it and none of them attempted a quantitative analysis about the degree of relatedness between the chosen domains. The lack of empirical analyses about relatedness in S3 documents can be attributed mainly to two reasons: on the one hand, the absence of a consolidated methodology to deal with the issue; on the other hand, the difficulty in performing quantitative measures, given for example that technological domains are identified using the natural language instead of a codified classification system. To overcome the latter problem, we have recoded the technological domains chosen by regions using the IPC system, looking at the most detailed description of such domains as indicated in the documents officially approved. The use of IPC codes provides an effective way of mapping the specialisation domains selected by regions. Moreover, it allowed us to perform a quantitative analysis aimed at assessing to what extent regions have been able to choose technological domains that show a high degree of relatedness. The measure of relatedness between technological domains is based on the revealed association methodology (Hidalgo et al., 2007). Based on the application of this methodology and using different ways to define the degree of relatedness, we measure to what extent the choices made by regions are better than a random choice. By applying this methodology, we evaluate the choices made by Italian regions in designing their Smart Specialisation Strategy.

In general, the choices made by regions showed a higher degree of relatedness compared to a random choice, though with notable exceptions. These exceptions would have been avoidable if these measures had been available during the design phase of the S3 rather than ex-post. As expected, the degree of relatedness is higher with respect to a random choice when we consider technological proximity based on absolute specialisation rather than on relative specialisation. Moreover, we suggest that the former is better than the latter in providing a meaningful measure of the intensity of relatedness between the technological domains that are actually present in a specific region. We compare these measures of relatedness to the degree of relatedness between the technological domains in which Italian regions are actually specialised. The results of this analysis confirm that in choosing S3 specialisation domains regions paid more attention in selecting those in which they had an actual strength rather than selecting a set of domains with the aim to maximize the degree of relatedness between them. This does not come as a surprise given that only few regions mentioned the concept of relatedness between specialisation domains and none attempted to apply a specific methodology to measure it.

The main contributions of this paper are both at methodological and empirical level. On the methodological side we provide a measure of relatedness between technological domains and propose and indicator of the degree in which regional choices were able to exploit the potential relatedness between related domains. On the empirical side, we perform a first attempt of an evaluation exercise of the degree of relatedness of regional choices within S3. In fact, empirical evidence about S3 is mainly indirect and aimed at providing insights about its theoretical arguments (see e.g. Balland, Boschma, Crespo, & Rigby, 2018) more than looking at its design, implementation and effects. Moreover, our methodology provides an ex-post measure of relatedness that may help regions to be aware of their specialisation choices and to monitor the variations in the relatedness index over the programming period. Most likely, the heterogeneity in the design of S3 policy will affect policy effectiveness, if any. Therefore, it must be taken into account in the evaluation of the results of the policy.

It's worth to point out that in this paper we are looking at S3 without expressing any judgement or evaluation of its underlying principles, of which relatedness is one of the core elements. We believe that in the ex-post assessment of its effectiveness and its underlying principle it is important to take into consideration whether and to what extent regions have actually implemented the policy following the "prescribed" recommendations. To evaluate the effectiveness of a medicine it is not sufficient to look at whether the doctor prescribed it without considering whether the patient actually bought it and took it following the doctor recommendations (e.g. about how frequently to take it).

It's worth mentioning that the principle of relatedness or better its theoretical foundation, namely related variety, has recently been criticized. For example, a growing body of literature started to point out its limitations in regions characterized by lower-technology regimes (Cortinovis & van Oort, 2015; Hartog et al., 2012). Other recent empirical contributions suggest for example that lagging regions may benefit especially from unrelated variety instead (Castaldi, Frenken, & Los, 2015; Firgo & Mayerhofer, 2018). This evidence clearly raises doubts also about the effectiveness of

S3 if applied as a *one size fits all* policy and about its potential effectiveness for less-developed regions or those locked-in in traditional industries and medium or low technologies. Notwithstanding these critiques, relatedness remained one of the basic principles of S3 and it is important to measure to what extend this principle was actually followed in the design of the policy.

The use of patents to measure regional innovative capabilities and technological specialisation represents both a major strength and a limitation of the present work. Patents are probably the most widely used indicator of innovation since Griliches (1990), notwithstanding their well-known flaws. In regional economics literature, patents are probably the most diffused measure to characterize regional knowledge capabilities/spaces (see e.g. Kogler, Essletzbichler, Rigby, 2016), to represent relatedness between knowledge domains (see e.g. Quatraro, 2010) and ultimately to provide insights and policy suggestions related to S3 (see e.g. Balland et al., 2018; Montresor & Quatraro, 2017). However, using patents as a measure of innovative activity and regional capabilities may limits the scope of our analysis while giving a strong emphasis on the technological aspects of S3; this fails to capture "social, organisational, market and service innovation, or practice-based innovation" and underrepresents sectorial specialisations in traditional sectors like for example tourism, agri-food, culture/heritage, wellness. For this reason, our methodology leads to an analysis of *patentable* technologies and formally codified knowledge, providing figures only about a portion (albeit probably the most relevant) of the whole "regional innovation landscape". However, we believe that, while being a partial representation, our analysis is not giving distorted evidence about regional technological oriented R&D activities. Moreover, a strong emphasis on R&D activities and outputs is evident not only in S3 guidelines but also when looking at the technological domains as indicated by regions in designing (and communicating) their S3 strategy. In fact, only a tiny fraction of those domains refers to non-patentable knowledge and only a small fraction of them is characterized by a low technological content. Finally, patents were chosen by EU and regional authorities as one of the most important indicators to monitor of the effectiveness of the S3.

As for any analysis based on past data, our measures are based on the world *'as it is'* (or better, *'as it was'*) rather than on what it is *'to be'*: i.e. regions may have chosen domains in which they are not yet specialised in but in which they are willing to specialise, i.e. where they see a 'potential' for specialisation.

Notwithstanding the above limitations, we believe that the methodology we propose is useful in providing a quantitative measure of the degree of relatedness between the technological domains chosen by regions. As such, it can help regions to be more aware of their choices and to monitor the implementation and the results of S3.

Moreover, this methodology could also be applied to measure the potential relations between technological domains in different regions. The latter was another aspect recommended in the S3 guide but generally overlooked by regions.

Finally, two policy messages are worth being pointed out. First, we believe that turning an academic argument into the implementation of a policy has been extremely difficult for regional authorities. This is because the concept of relatedness is something difficult to measure (and sometimes even to understand) even for academics. In fact, there is no single unanimous methodology to measure it and, in our opinion, there is also a lack of understanding of the "inner" mechanisms, i.e. "below the regional level". Second, another crucial point from a policy perspective is that the effectiveness of S3 should not be given as granted. Specifically, the effectiveness of this policy is likely to be extremely influenced by its actual design and implementation. Our evidence is that there is great heterogeneity in its implementation (at least in Italy) and that most of its principles were far too difficult to turn into actual plans (and subsequent actions) for regional authorities.

ORCID

Diego D'Adda https://orcid.org/0000-0002-1683-1787
Donato Iacobucci https://orcid.org/0000-0001-8463-1106

REFERENCES

Asheim, B. T., Boschma, R., & Cooke, P. (2011). Constructing Regional Advantage: Platform Policies Based on Related Variety and Differentiated Knowledge Bases. *Regional Studies*, *45*, 893–904.

Asheim, B. T. and Grillitsch, M. (2015). Smart Specialisation: Sources for New Path Development in a Peripheral Manufacturing Region, *Papers in Innovation Studies No. 2015/11*, *CIRCEL; Lund Univeristy*, 1–27.

Balland, P.-A., Boschma, R., Crespo, J. and Rigby, D. L. (2018). 'Smart Specialization Policy in the European Union: Relatedness, Knowledge Complexity and Regional Diversification', *Regional Studies*, 1–17.

Balland, P. A., & Rigby, D. (2017). The Geography of Complex Knowledge. *Economic Geography*, *93*, 1–23.

Boschma, R. (2014). Constructing Regional Advantage and Smart Specialization. *Scienze Regionali*, *Italian Journal of Regional Sciences*, *13*, 51–68.

Boschma, R., & Frenken, K. (2011a). Technolgical Relatedness, Related Variety and Economic Geography. In P. Cooke, B. T. Asheim, R. Boschma, R. Martin, D. Schwartz, & F. Todtling (Eds.), *Handbook of Regional Innovation and Growth* (pp. 187–197). Cheltenam (UK): Edward Elgar.

Boschma, R., & Frenken, K. (2011b). Technological Relatedness and Regional Branching. In H. Bathelt, M. Feldman, & D. F. Kogler (Eds.), *Beyond Territory: Dynamic Geographies of Knowledge Creation, Diffusion and Innovation* (pp. 64–81). London: Routledge.

Boschma, R. and Gianelle, C. (2014). Regional Branching and Smart Specialisation Policy, *S3 Policy Brief Series*.

Boschma, R., & Iammarino, S. (2009). Related Variety, Trade Linkages, and Regional Growth in Italy. *Economic Geography*, *85*, 289–311.

Boschma, R., Minondo, A., & Navarro, M. (2010). *Related Variety and Regional Growth in Spain*. Urban & Regional Research Centre Utrecht: Utrecht.

Breschi, S., Lissoni, F., & Malerba, F. (2003). Knowledge-Relatedness in Firm Technological Diversification. *Research Policy*, *32*, 69–87.

Capello, R. (2013). Knowledge, Innovation, and Regional Performance: Toward Smart Innovation Policies. *Growth and Change*, *44*, 185–194.

Capello, R., & Kroll, H. (2016). From Theory to Practice in Smart Specialization Strategy: Emerging Limits and Possible Future Trajectories. *European Planning Studies*, *24*, 1393–1406.

Castaldi, C., Frenken, K., & Los, B. (2015). Related Variety, Unrelated Variety and Technological Breakthroughs: An Analysis of US State-Level Patenting. *Regional Studies*, *49*, 767–781.

Cortinovis, N., & van Oort, F. (2015). Variety, Economic Growth and Knowledge Intensity of European Regions: A Spatial Panel Analysis. *The Annals of Regional Science*, *55*, 7–32.

D'Adda, D., Guzzini, E., Iacobucci, D., & Palloni, R. (2018). Is Smart Specialisation Strategy Coherent with Regional Innovative Capabilities? *Regional Studies*, 1–13.

Dalum, B., Laursen, K., & Villumsen, G. (1998). Structural Change in OECD Export Specialisation Patterns: de-specialisation and "stickiness". *International Review of Applied Economics*, *12*(3), 423–443.

Duranton, G., & Puga, D. (2001). Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products. *American Economic Review*, *91*, 1454–1477.

Firgo, M., & Mayerhofer, P. (2018). (Un)Related Variety and Employment Growth at the Sub-regional Level. *Papers in Regional Science*, *97*, 519–547.

Foray, D. (2015). *Smart Specialisation. Opportunities and Challenges for Regional Innovation Policy*. London: Routledge.

Foray, D., David, P. A. and Hall, B. (2009). Smart Specialisation – The Concept, Brussels, 9.

Foray, D., David, P. A. and Hall, B. H. (2011). Smart Specialisation - From Academic Idea to Political Instrument, the Surprising Career of a Concept and the Difficulties Involved in Its Implementation.

Foray, D., Goddard, J., Beldarrain, X. G., Landabaso, M., McCann, P., Morgan, K., Nauwelaers, C., Ortega-Argilés, R., & Mulatero, F. (2012). Guide to Research and Innovation Strategies for Smart Specialisations (RIS 3), Brussels.

Foray, D. and Goenega, X. (2013). The Goals of Smart Specialisation.

Frenken, K., Van Oort, F., & Verburg, T. (2007). Related Variety, Unrelated Variety and Regional Economic Growth. *Regional Studies*, *41*, 685–697.

Griliches, Z. (1990). Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, *28*, 1661–1707.

Grillitsch, M., Asheim, B., & Trippl, M. (2018). Unrelated Knowledge Combinations: The Unexplored Potential for Regional Industrial Path Development. *Cambridge Journal of Regions, Economy and Society*, *11*, 257–274.

Hartog, M., Boschma, R., & Sotarauta, M. (2012). The Impact of Related Variety on Regional Employment Growth in Finland 1993–2006: High- Tech versus Medium/Low-Tech. *Industry and Innovation*, *19*, 459–476.

Hidalgo, C. A., Klinger, B., Barabasi, A.-L., & Hausmann, R. (2007). The Product Space Conditions the Development of Nations. *Science*, *317*, 482–487. https://doi.org/10.1126/science.1144581

Iacobucci, D. (2014). Designing and Implementing a Smart Specialisation Strategy at Regional Level: Some Open Questions. *Scienze Regionali, Italian Journal of Regional Science*, *13*, 107–126.

Iacobucci, D., & Guzzini, E. (2016). Relatedness and Connectivity in Technological Domains: The "Missing Links" in S3 Design and Implementation. *European Planning Studies*, *24*, 1511–1526.

Jacobs, J. (1969). *The Economy of Cities*. New York: Random House.

Kogler, D. F., Essletzbichler, J., & Rigby, D. L. (2016). The evolution of specialization in the EU15 knowledge space. *Journal of Economic Geography*, *17*(2). https://doi.org/10.1093/jeg/lbw024

Kogler, D. F., Essletzbichler, J., & Rigby, D. L. (2017). The Evolution of Specialization in the EU15 Knowledge Space. *Journal of Economic Geography*, *17*, 345–373.

Lambooy, J. G., & Boschma, R. (2001). Evolutionary Economics and Regional Policy. *Annals of Regional Science*, *35*, 113–132.

McCann, P., & Ortega-Argilés, R. (2013). Modern Regional Innovation Policy. *Cambridge Journal of Regions, Economy and Society*, *6*, 187–216.

McCann, P., & Ortega-Argilés, R. (2015). Smart Specialization, Regional Growth and Applications to European Union Cohesion Policy. *Regional Studies*, *49*, 1291–1302.

Montresor, S., & Quatraro, F. (2017). Regional Branching and Key Enabling Technologies: Evidence from European Patent Data. *Economic Geography*, *93*, 367–396.

Neffke, F., & Henning, M. (2013). Skill Relatedness and Firm Diversification. *Strategic Management Journal*, *34*, 297–316.

Neffke, F., Henning, M., & Boschma, R. (2011). How Do Regions Diversify over Time? Industry Relatedness and the Development of New Growth Paths in Regions. *Economic Geography*, *87*, 237–265.

Quatraro, F. (2010). Knowledge Coherence, Variety and Economic Growth: Manufacturing Evidence from Italian Regions. *Research Policy*, *39*, 1289–1302.

How to cite this article: D'Adda D, Iacobucci D, Palloni R. Relatedness in the implementation of Smart Specialisation Strategy: a first empirical assessment. *Pap Reg Sci*. 2020;99:405–425. https://doi.org/10.1111/pirs.12492

APPENDIX A

PROXIMITY À LA HIDALGO BASED ON RELATIVE SPECIALIZATION MEASURES

Following Hidalgo et al. (2007), we compute the proximity matrix, representing the degree of proximity between pairs of IPC classes, that is, a measure of the probability that a region is specialized in a specific pair of IPC classes. Proximity matrix is derived from the actual specialization of (272) European regions, as derived by their actual patenting activity.

The values of proximity vary between 0 and 0.684, with mean equal 0.270 and median equal to 0.263. We report below the histogram (see Figure A1) of these values and some descriptives (see Table A1).

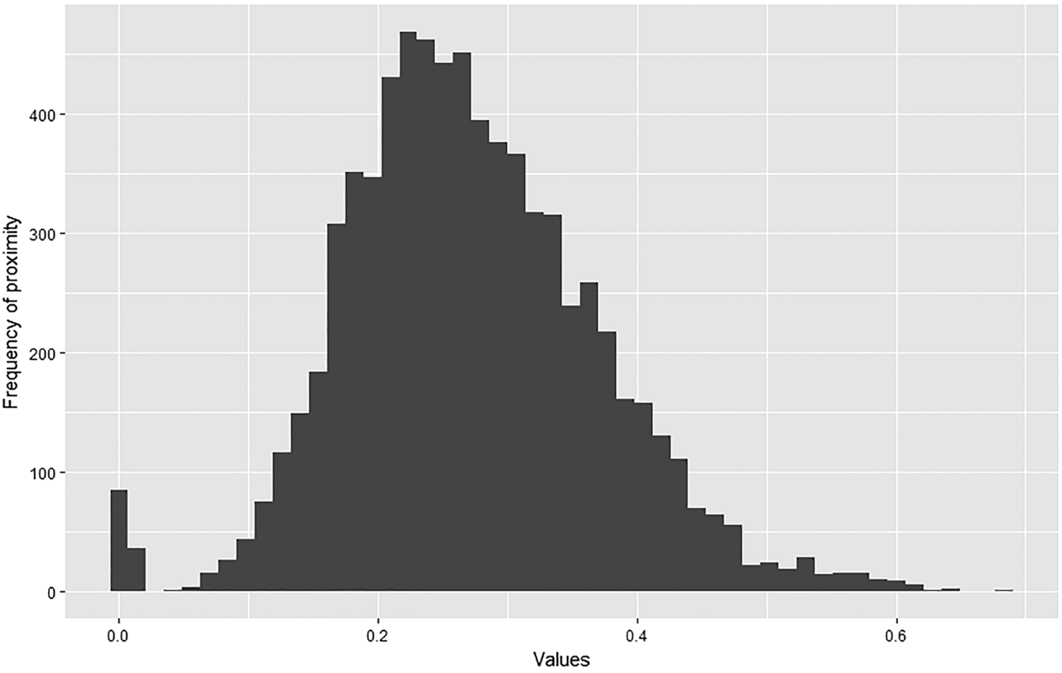

FIGURE A1 – Distribution of the proximity (between all pairs of IPC codes) à la Hidalgo based on relative specialization measures

TABLE A1 Summary statistics of the proximity (between all pairs of IPC codes) *à la* Hidalgo based on relative specialization measures

Min	1st quartile	Median	Mean	3rd quartile	Max
0	0.2051	0.2632	0.2708	0.3306	0.6839

To give an example, there are some IPC classes with a high degree of proximity (above 0.48, please note that the third quartile is equal to 0.33) related to the technologies applied to metals and mechanics:

Appendix A. PROXIMITY À LA HIDALGO BASED ON RELATIVE SPECIALIZATION MEASURES

	B21	B22	B23	
B21	NA	0.481481	0.607595	MECHANICAL METAL-WORKING WITHOUT ESSENTIALLY REMOVING MATERIAL; PUNCHING METAL
B22	0.481481	NA	0.530864	CASTING; POWDER METALLURGY
B23	0.607595	0.530864	NA	MACHINE TOOLS; METAL-WORKING NOT OTHERWISE PROVIDED FOR

On the contrary, an example of two IPC classes with a low proximity (close to 0) is for example aircraft (B64) and textiles (D06) technologies.

Appendix A. PROXIMITY À LA HIDALGO BASED ON RELATIVE SPECIALIZATION MEASURES

	B64	D06	
B64	NA	0.068493	AIRCRAFT; AVIATION; COSMONAUTICS
D06	0.068493	NA	TREATMENT OF TEXTILES OR THE LIKE; LAUNDERING; FLEXIBLE MATERIALS NOT OTHERWISE PROVIDED FOR