"With no siblings, but with home": Single Patent-code Inventions (SPIs) and regional technological diversification

Claudio Cozza[♣]

(claudio.cozza@uniparthenope.it)

Sandro Montresor[♦]

(sandro.montresor@gssi.it)

Alessandro Palma[♦]

(alessandro.palma@gssi.it)

Gianluca Tarasconi[♠]

(gt@ipQuants.com)

Abstract

We investigate the extent to which the "siloed" nature of regions' knowledge base affects their technological diversification. Drawing on the recombinant innovation theory, we maintain that a more siloed knowledge base for technologies in regions make more difficult for them to gain the relative specialisation over time, and that this effect is more pronounced for more cognitively unrelated new technologies. Using EPO patent data for EU-28 (NUTS2) regions, we proxy the siloed extent of regional technological knowledge with the incidence of Single Patent-code Inventions (SPIs) and test these hypotheses over the 1986-2017 period. Consistently with them, we find that the local SPI incidence in technologies negatively correlates with their regional entry for an average level of relatedness, and that the correlation attenuates when such an average increases. However, results are conditional on very high levels of relatedness, for which the SPI incidence turns to enable, rather than contrasting technological diversification. Policy implications are drawn accordingly.

Key Words: Single patent-code inventions; innovation geography; relatedness; regional technological diversification.
JEL codes: R11, R58, O31, O33

[♣] DiSEGIM Department, University of Naples "Parthenope" (IT).
[♦] Gran Sasso Science Institute, L'Aquila (IT).
[♦] Gran Sasso Science Institute, L'Aquila (IT).
[♠] ipQuants AG, Schaffhausen (CH).

1. Introduction

Local specialisations, denoted by those economic activities in which regions specialise at a certain moment in time, are exposed to a constant pressure of renewal. As the economic effects of the Covid-19 pandemic and of the Russia-Ukraine war have shown to an unprecedent scale, an important part of this pressure comes from external shocks. These might even force regions to restart their business operations from scratch in an uncharted scenario. The push to change is however also and above all endogenous, due to the stagnation and decline of the activities in which regions specialise, following the unfolding of techno-economic evolution and competition (Xiao et al., 2018). The incessant pression of these forces has led to identify in regional diversification a crucial leverage of competitiveness and structural change and has put the analysis of its determinants on the top of both research and policy agendas (Neffke et al., 2011).

Among the different approaches adopted in its analysis, the so-called "relatedness" approach has shown that regional diversification mainly occurs as the "branching" of pre-existing activities, to which newly developed ones are cognitively related (Neffke et al. 2011; Boschma and Gianelle, 2014, Hidalgo et al., 2018). Despite important nuances and exceptions (Zhu et al., 2017), regions predominantly enter new activities that share similar capabilities to pre-existing ones, confirming the theoretical predictions about the role of path-dependence, local search and routinised behaviors in evolutionary economic geography (Boschma and Frenken, 2006).

The relatedness approach has been applied also and above all to the analysis of regional technological diversification, amounting to the regions' capacity of innovating and entering previously unmastered technological domains. This analysis is carried out by extending to the regional realm the recombinant theory of firms' innovations (Weitzman, 1998) and claiming that, through the meso-aggregation of the latter, new regional technologies mainly emerge through novel recombinations of pre-existing ones (Balland, 2016). Indeed, within this theoretical framework, relatedness captures the cognitive proximity between acquired and pre-existing technologies and appears a crucial facilitator of the recombinantion of extant local ideas into new ones (Boschma et al., 2015; Rigby, 2015; Kogler et al., 2013). On this basis, a normative prediction has been derived, encouraging regions to develop smart specialization strategies (S3) by supporting related (technological) diversification (Boschma and Giannelle, 2014).

Despite this ample stream of literature, we claim that an important aspect has so far been neglected in grounding the relatedness approach at stake in the recombinant innovation theory. While a due focus has been placed on the role of cognitive proximity *between* technologies in the recombination, little if no attention has been paid to the range of knowledge domains that regions have available to recombine *within* specific technologies. The knowledge underpinning some technologies can be pretty "simple", while that of other ones more "complex", for example, in terms of multi-disciplinarity and interconnectedness, and this crucially affects the way ideas can be combined and recombined with respect to them. This aspect has been recognised as very relevant in micro and sectoral innovation studies, particularly in those about "technological regimes" (Winter, 1984; Malerba and Orsenigo, 1996), and its neglect in regional ones about technological diversification appears an unfortunate gap.

In filling this gap, we argue that the relatedness argument and results could be affected by the extent to which the knowledge that underpins regional technologies – i.e., their knowledge base (Malerba and Orsenigo, 1996) – appears marked by unique, rather than multiple, domains or, metaphorically, by "knowledge silos". We argue that, with respect to technologies whose regional knowledge is (more) "siloed", diversification could be harder to implement, given the factual gap (if not even absence) of knowledge-combination opportunities to exploit locally, on which technological diversification relies. On the same basis, we also argue that the disadvantages of targeting more siloed technologies in knowledge terms could be attenuated by their higher relatedness to existing ones. Indeed, the knowledge combinatorial gaps entailed by a more siloed technology for the sake of diversification, could be compensated by the benefits of pursuing a technology that draws on more similar capabilities to existing ones.

To test these arguments, we propose to proxy the siloed nature of local technologies by measuring the extent to which the relative inventive activities occur only in one knowledge domain. Alluding to the "codes" (in generic terms) attributed to patent applications (by the applicants and/or the examiners) with respect to existing patent classification systems (e.g., IPC or CPC), this can be captured by the incidence of "Single Patent-code Inventions" (SPIs) in regional technologies. We then plug this region-technology-specific SPI indicator in a standard econometric model of regional technological diversification based on relatedness – i.e., a regional technological branching model – and address the relationship between the two with respect to the EU28 NUTS2 regions over the period 1986-2017.

To start with, our analysis shows that, with respect to the "global" knowledge space (Balland, 2016), the incidence of SPIs is far from negligible, regardless of the IPC aggregation level and of the considered geographical area, thus deserving investigation also and above all in the analysis of regional technological branching. As we did expect, conditionally on an average level of relatedness, the incidence of local SPIs by technology shows a significant negative relationship with the regional capacity to specialize in it *ex-novo*: in other word, a higher SPIs' incidence and a more siloed knowledge base of technologies, actually hampers the regional diversification in them. Furthermore, as we also expected, a higher average level of relatedness negatively moderates the impact that the incidence of SPI in a regional technology has on the region's capacity to diversify in it. However, results appear conditional on, and differ with, the actual level of relatedness of new technological entries. When this is quite high, a siloed knowledge base even switches the sign of its relationship with technological diversification into a positive one.

These results provide an important additional rationale for recommending regions to pursue smart specialization strategies in research and innovation (S3). Indeed, implementing this policy in stringent terms, by targeting highly related technologies, makes of this policy an effective remedy to reverse the difficulties in diversifying faced by regions with a more siloed knowledge base. Accordingly, the measurement of this last feature of the regional innovation system appears a crucial aspect to consider by regional policy makers involved in the implementation of S3.

The rest of the paper is structure as follows. Section 2 illustrates the conceptual and empirical background of our idea of siloed regional knowledge base and of the SPI proxy that we propose. Section 3 illustrates the empirical application and Section 4 the relative results. Section 5 concludes.

2. Background studies

2.1 Theoretical background and research hypotheses

The geography of innovation reveals an uneven regional distribution of new knowledge production and technological novelty. Regions present heterogeneous knowledge bases (Asheim, 2007), marked by different combinations of knowledge typologies (e.g. synthetic, analytical, and symbolic) and by networks of knowledge domains with diverse structures (Asheim and Coenen, 2005). Following the Schumpeterian account of innovation as '*Neue Kombinationen*' of ideas (Becker et al., 2012; Weitzman, 1998; Evenson and Kislev, 1976), this heterogeneity in regional knowledge bases has been taken to explain the emergence of different patterns of regional innovations as emerging from the re-combinations of their constitutive ideas. In the revival of this Schumpeterian account by the so-called "relatedness approach" (Balland, 2016), the regional knowledge base has been argued and shown to affect the regions' capacity to diversify their technologies over time. In brief, following this approach, regions enter new technological fields that are cognitively proximate to pre-existing ones, as they require similar capabilities (Boschma et al., 2015; Rigby, 2015; Kogler et al., 2013).

In developing this argument, the extent to which the knowledge base of regions does actually provide opportunities of innovative re-combinations has been mainly interpreted in Jacobsian terms (1969) and accounted by its "variety": more precisely, by the variants along which the variety concept has been lately disentangled, e.g. related and unrelated variety (Frenken et al., 2007). In brief, a variegated technological knowledge in the region has been claimed and found responsible of spillovers and cross-fertilisation of ideas, which can lead regions to innovate even in a radical manner (Castaldi et al., 2015; Mewes, 2019; Berkes and Gaetani, 2020). The relatedness approach to diversification in fact draws on the benefits of knowledge variety and, as a step forward, rather than opposing it to the Marshallian idea of specialization, it synthesizes the two in that of "diversified specialization" (Farhauer and Kroll, 2012).

While the previous kind of variety of the knowledge base is for sure relevant, the extent to which recombinant innovations occur and lead regions to develop new technologies can be affected by a different feature of the regional knowledge base, which could be metaphorically denoted as its cognitively "siloed" nature. Rather than to the distribution of regional inventive efforts/outcomes across different knowledge domains and technologies, to which the idea of variety refers, that of siloed knowledge base that we propose to investigate, looks at the extent to which technologies draw on single knowledge domains, evoking the idea of technological trajectories that appear like "silos".

Looking at technological change according to a knowledge-based theory of production (Rosenberg, 1976), the array of knowledge domains that underpin a focal technology, represents an important element for characterizing the cognitive environment in which the problem-posing and problem-solving activities leading to innovations take place: in Nelson and Winter (1982)'s terminology, of the relevant "technological regime". As evolutionary economics has shown since long (Dosi, 1982; Winter 1984; Malerba and Orsenigo 1996; Breschi et al., 2000), technological regimes differ among them along two connected levels of analysis: i) the conditions with which innovation processes take place in the regime;[1] ii) the characteristics of the knowledge inputs and outputs of the same innovation processes, typically represented by the tacitness, observability, complexity, and systemic

[1] These conditions refer to technological opportunities (e.g., high rather than low probability (risk) to innovate (non-innovate)), appropriability (e.g., ease vs. difficulty of protection against imitation), and of cumulativeness (e.g., high vs. low path-dependence and innovation persistence).

nature of knowledge (Winter, 1987). In extreme nuthsell, tacitness refers to the extent to which technological knowledge can be articulated into symbols and codes, like in patent documents, and observability to that it can be fully disclosed and inspected, like with the acquisition, license and/or expiration of patents. As for complexity, low-complex technological regimes distinguish from high-complex ones for the small, rather that high number of knowledge items and fields that mark their knowledge base (Breschi et al., 2000). Similarly, high-systemic and low-systemic technological regimes differ for knowledge items that are largely interconnected rather than independent in the development of the focal technology (ibid.).

Focusing on these last two features, and contextualizing them in a regional framework, the incidence of “knowledge silos” in the regional knowledge base can be associated to the pervasive presence of technological regimes that are “simple” and marked by a low or no-complex / systemic nature. Such an incidence has important implications for the regional capacity of specializing in them ex-novo, that is, of diversifying their technological repertoire. These implications descend from the fact that, as we have noticed above, technological diversification mainly draws on the regions’ capacity to explore and envisage new combinations among different knowledge items of a certain technological domain, which in the case of a siloed knowledge for it are drastically reduced or even absent. Indeed, a siloed regional knowledge for a certain technology, being accompanied by a negligible or even nil extent of multi-disciplinarity, dampens the set of knowledge linkages that can be explored for the region to diversify in it. As we will argue in the following, a typical case of siloed knowledge is that of a technology in which regional inventive activities lead to patents that refer to a single “code”, among those in which technological knowledge is classified following available international schemes at patent offices (like IPC and CPC). As we will also see in the following, this is not an extreme event at all, especially in some technologies like medical and digital communication, in which the share of what we will call “Single Patent-code Inventions” (SPIs), is appreciable. In front of such a kind of patents, it could be claimed that scope of possible combinations could be stimulated by the use that inventions within a certain technological domain makes of the inventive knowledge obtained in other domains, as it is typically reflected by the relative citation patterns. Citations surely convey knowledge links, which could counteract the silo-effect entailed by single patent-code inventive activities. However, because of their nature of knowledge spillovers, citations might represent less actual and workable knowledge links than those represented by the simultaneous presence of different knowledge items in the same invention (co-occurrence) (Balland, 2016). Furthermore, for simply computational reasons, the extent to which citations unfold and interconnect the knowledge base, decreases with the extent to which inventions are based on single knowledge domains.

Based on the previous arguments, we expect that a high incidence of inventive activities in single knowledge domains renders the relative regional technological knowledge "thin" in terms of combination opportunities and makes technological diversification harder to occur. Using our idea of the siloed nature of the knowledge underlaying a certain technology, our first research hypothesis is thus the following:

Hp1: *The more siloed is the regional knowledge base of technologies, the lower is the regional capacity to diversify in them*.

A siloed kind of knowledge for regional technologies does also interfere with the role that relatedness has in influencing technological diversification. As we have recalled above, regions have been pervasively found to entry into new technologies that are cognitively related to their pre-existing ones; targeting technologies marked by higher relatedness has accordingly become a policy prescription in the S3 framework. A higher level of relatedness facilitates technological diversification by enabling regions to control new technologies, which draw on capabilities similar to those underlying the existing ones in the knowledge base (Balland, 2016). This similarity in the required capabilities represents an important diversification enabler, which could make the lack of knowledge combinatorial opportunities of a siloed technology less hampering for it. In other words, a new technology with a certain degree of siloed knowledge is less difficult to master by regions in which it is closer to the existing knowledge base. As Castellani et al. (2022) have found by investigating the (green) technology diversification role of inward FDIs in a region, the level of relatedness between new and existing technologies represents an important moderating, and even conditioning, factor of the effect exerted by a focal diversification enabler. In our case, the expected moderation effect is negative and lead us to put forward the following hypothesis:

Hp2: *The higher the relatedness of new technologies, the lower is the impact of their siloed knowledge base on the regional capacity to diversify*.

2.2 Empirical background

The idea of a siloed regional knowledge (base), and its role in the conceptual framework of the relatedness approach, can be better grasped by looking at the way this approach has been implemented in empirical research, that is, using patent data (Acs et al., 2012). Following this stream of literature, the technological codes through which patents are classified at patent offices can be taken to identify the so-called ‘knowledge space’: a network of technological fields, whose structure affect the dynamics of regional innovation and industries, thus inspiring regional policies for their ‘smart’ development (Rigby, 2015; Balland et al., 2018).

As the idea of knowledge combination has turned out pivotal in innovation geography, searching for combinations among patent codes has become an essential tool to map the knowledge space and, by locating regions within it, the structure of their knowledge base. More precisely, knowledge combinations have been mapped either by looking at the co-occurrence of the different codes with respect to which patents claim to have brought novelty, or by following the citations that patents make among their relative codes (usually the primary ones). In both kinds of search for technological relatedness, and particularly in that based on the frequency with which two technological codes appear on the same patent (see Balland et al., 2018), the number of patents that present more than one code reveals of course decisive. Indeed, those patents to which only one patent code is assigned do not make, by definition, combinations between components and/or principles, irrespectively from their being novel or not.

As we will see in the following section, patents of this kind, which we consider denoting Single Patent-code Inventions (SPIs), can be retained to provide an interesting proxy of the siloed knowledge base of a certain regional technology addressed in the previous sub-section. In the light of that, they are an important typology of patents. However, in the studies on the identification and geographical distribution of technological novelty by this knowledge re-combination (e.g., Verhoeven et al., 2016; Mewes, 2019), these SPIs are usually ruled out and/or retained uninformative of the knowledge base of the regions where they locate[2].

The neglect of SPIs in the regional distribution of (radical) innovation is however unfortunate also from an empirical point of view, as their incidence is far from exiguous. Referring to the PATSTAT dataset (Spring 2022 version, consulted on 7/2/2023), and to the International Patent Classification (IPC), it emerges that, at its highest level of disaggregation (full IPC or “subgroups”, amounting to

[2] For example, in their recent study on technological novelty, Verhoeven et al. (2016, p. 711) maintain that: “When only 1 IPC group code is assigned, our […] indicator of technological novelty is underdefined, as we fail to identify at the IPC group level the recombination of components/principles it is making. In our analyses, we treat these cases as missing values when we analyze the […] indicator. […] Alternatively, we could consider the patents with only 1 IPC group code as not making any substantial recombination that cross IPC groups and therefore also not having any novel recombination and thus scoring zero on [our indicator]”.

70,191 codes), around 20% of the world patent applications are SPIs. Such a share holds for all main patent offices (EPO, USPTO, JPO and WIPO) over the last decades and it even overcomes the share of 30% in entering the 2000s.

The overall relevance of SPIs across all patent offices is confirmed, and obviously increases, when looking at a more aggregated level of their classification codes. Figure 1 reports the share of SPIs at the IPC4[3] or "subclasses" (647 in total) level for year 2017, that is the last one of our empirical analysis. In 2017 the SPI share is around the 50% across the board and, although this share for the USPTO in year 2017 is just above 40%, it is closer or even higher than 50% in other years of the time-series (e.g., 50.3% in 2012 or 52.1% in 2008).

Figure 1 – Share of SPIs on total patent applications, at the IPC4 level, by patent office (2017)

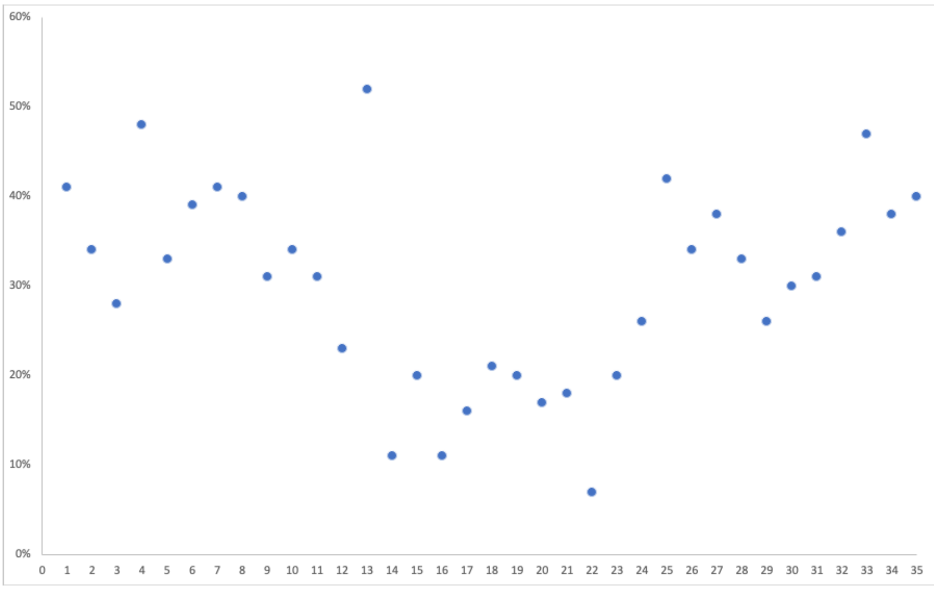
Source: own elaboration on PATSTAT, Spring 2022 version

Overall, the evidence about the relevance of SPIs does not change that much if we refer to the applications at a single patent office, like the European Patent Office (EPO), as we do hereafter and in our empirical application. The SPI phenomenon is however heterogeneous across different kinds of technologies, as Figure 2 shows using the aggregation of IPC subclasses into the 35 technological fields proposed by Schmoch (2008). In 2017, for all the considered 35 fields, the share of SPIs is at least 10%, with the only exception of field 22 (Micro-structure and nano-technology). The share is

[3] At this level of disaggregation, a focal patent is retained SPI even if it includes more IPC codes (full IPC, that is the most disaggregated ones as in Figure 1), but they all belong to the same IPC 4 digit "subclass".

the highest and close to 50% for medical technology (13) and digital communication (4), while the lowest incidence of SPIs (lower than 20%) can be found, as expected, given their "complexity", in fields as pharmaceuticals (16), organic fine chemistry (14), macromolecular chemistry, polymers (17), surface technology, coating (21), and materials, metallurgy (20).

Figure 2 – Share of SPIs on total patent applications at EPO by IPC field: 2017



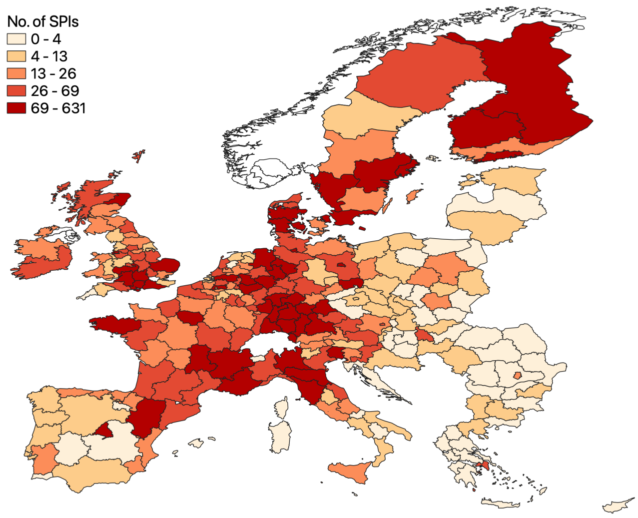
Source: own elaboration on PATSTAT, Spring 2022 version

The heterogenous picture revealed by Figure 2 suggests us to investigate the incidence of SPIs by carefully retaining their technological specificity: that is, at the technology level. Further suggestions emerge by looking at the geographical distribution of SPIs across European regions. A definitively non-negligible number of regions host the development of technologies whose regime is, not only "non-complex", but as we said, "the simplest". Quite interestingly, Figure 3 reveals that several EU regions with the highest number of SPIs are also peripheral regions, hinting at the location of the simplest technological regimes precisely in these lagging behind territories. However, the map also shows strong heterogeneity within each country – also the most advanced ones – and among leading/peripheral regions across countries. Unlike other innovation variables, which mostly map in the classical core-periphery gradient of European regions (see, for instance, Marsan & Maguire, 2011; Capello & Lenzi, 2013; De Noni et al., 2018), we detect an only partial overlapping between SPIs and lagging-behind regions and countries, which for sure deserves attention in future research.

The distribution of SPIs across technologies and regions represents the starting point of our empirical analysis of technological diversification, to which we move in the next Section.

Figure 3 – Number of SPIs by European NUTS2 regions (2014-2017)

Commentato [CC1]: Viene molto diversa da quella precedente perché in valore assoluto, mentre l'altra era in percentuale… non so se ha senso così, perché ovviamente vengono più scure le regioni che brevettano di più



3. Empirical application

We test our research hypotheses (see Section 2) using a brand-new dataset that we have built up with respect to 264 NUTS2 regions of the EU28 over the period 1986-2017.[4] This dataset was obtained by inspecting regional patent data, extracted from raw patent applications to the EPO in the EPO PATSTAT database (Spring 2022 version), and by merging them with other regional data from the EUROSTAT and the European Regional Database maintained by Cambridge Econometrics.

[4] The upper boundary of this temporal window is due to data availability in connection with data truncation with respect to the most recent patents, while the lower boundary is consistent with that of previous studies about regional technological diversification for the EU (see, for example, Balland et al., 2018). As we refer to pre-Brexit period, our analysis considers the EU28. However, given the change in NUTS classification for some EU countries, minor adjustments have been needed: Greek regions EL41 and EL62 have been excluded; values for UKI1 region result from the average of UKI3 and UKI4 regions; values for UKI2 region result from the average of UKI5, UKI6 and UKI7 regions.

Referring to the extant literature, patent data have been fractionally regionalised using the inventor address and technologically classified according to the International Patent Classification (IPC), at the IPC4 or "subclass" level,[5] obtaining a set of 638 IPC subclass codes.

3.1 Dependent variable

Using the analytical framework of regional technological branching (Tanner, 2014), our focal dependent variable is represented by the "entry" of a new technology in the regional knowledge base. Consistently with this literature, such an entry measures the regional capacity of getting a new technological specialisation at time *t*: that is, a specialisation in a certain technology, *c*, which region *r* did not have at *t* – *k*. Measuring this technological specialisation through an index of Revealed Technological Advantages (RTA_{rct}), amounting to a Balassa indicator of Revealed Comparative Advantages obtained with patent instead of export data, our dependent variable is a dummy indicator defined as follows:

$$NewRTA_{rct} = 1,\ \ if\ \ RTA_{rct} \geq 1\ and\ 0 \leq RTA_{rct-k} < 1 \qquad (1)$$
$$NewRTA_{rct} = 0, otherwise$$

where RTA_{rc} is usually defined as $^{Q}/_{Q_r}$ according to the following definitions and positions:

$Q = \frac{n_c}{n}$; $Q_r = \frac{n_{rc}}{n_r}$

n: total number of EU-28 patents;

n_c: total number of EU patents having IPC4 code c;

n_r: total number of patents in region *r*;

n_{rc}: number of patents in region *r* having IPC4 code *c.*

As usual, the presence (absence) of a regional specialization in technology *c* is signalled by a share of regional patents in the same technology, which is higher (lower) than the regional share of total patents, that is, by $\frac{Q}{Q_r} \geq 1$ $(0 \leq \frac{Q}{Q_r} \leq 1)$. The entry of the same technology *c* in the knowledge base of region *r* is denoted by the shift from the absence to the acquisition of the same specialization from *t* – *k* to *t*: using *k* = 1 as a benchmark. To smooth the typical erratic trend of patent data over

[5] At this level of disaggregation, a focal patent is retained SPI even if it includes more IPC codes (full IPC, that is the most disaggregated ones as in Figure 1), but they all belong to the same IPC 4 digit "subclass".

Commentato [CC2]: PERCHè SOPRA SONO DI PIU?].
Ho leggermente modificato sopra, evidenziando che a questo livello di analisi stiamo guardando solo i brevetti chiesti all'EPO. In questo sotto-insieme, le classi sono 638 mentre nei dati sottostanti alla figura 1 (in cui c'erano anche USPTO, JPO e WIPO) si arrivava a 647; nulla di strano.

INOLTRE, E Più IMPORTANTE, POSSIAMO DIRE CHE LIVELLI Più DISAGGREGATI DI IPC AVREBBERO CREATO QUALCHE PROBLEMA ECONOMETRICO? perché DIVERSAMENTE OCCORRE DARE UNA GIUSTIFICA, SE NO CI CHIEDONO UN CONTROLLO CON UN ALTRO LIVELLO
ALESSANDRO: LA DISAGGREGAZIONE DEGLI IPC NON E' MAI CAMBIATA DALL'INIZIO E CREDO/SPERO (MA QUESTO LO SAPETE MEGLIO VOI) CHE SEGUA LO STANDARD NELLA LETTERATURA! OVVIAMENTE SE CAMBIA LA GRANULARITA' CAMBIERANNO ANCHE I RISULTATI MA QUESTO VALE PER TUTTI I PAPER CHE UTILIZZANO PATENTS...UNICO MODO DI ADDRESSARE QUESTO CONCERN E' DIRE CHE SEGUIAMO LA LETTERATURA PRECEDENTE (CON LA SCUSA DI BENCHMARK RESULTS ETC..)

Commentato [CC3]: Scusate, forse mi sto impicciando il cervello: ma non dovrebbe essere Qr al numeratore e Q al denominatore?

time, we follow previous literature (e.g., Montresor and Quatraro, 2019) and define our dependent variable and the regressors by subdividing the reference period into 8 four-year sub-periods, the first one being 1986-1989 and the last one 2014-2017.

3.2 Focal regressors and controls

The first focal regressor of our analysis is an indicator of regional SPIs by technology, *c*, which proxies the siloed nature of the relative regional knowledge. Consistently with the extant literature on the topic, we do not claim for causality in inspecting the possible drivers of regional technological diversification, including this one, and just inspect for correlations. Still, in trying to attenuate endogeneity issues in terms of reverse causality, this indicator and the other regressors are lagged with respect to the dependent variable and defined at *t – 1*.[6]

The indicator that we propose, SPI_{rct-1}, is both technology and region specific, and measures the siloed nature of the regional knowledge of a certain technology in relative terms, with the share of mono-IPC patents within each focal IPC4 code. Netting out the scale of inventive activities, which would affect the simple count of mono-IPC patents within each IPC4 code, SPI_{rct-1} provides us with more reliable information about the extent to which a siloed kind of knowledge is characteristic of a certain technology *c*. Using the same notation as before, we define it as:

$$SPI_{rct-1} = Q_{rct-1} = \frac{n_{rc,t-1}^{mono}}{n_{rc,t-1}} \qquad (2)$$

In addition to the previous SPI indicator, another focal regressor of our analysis is represented by the relatedness of the new technology *c* in which the region specialises at *t*, to the technologies it has already specialised at *t – 1*. In particular, we follow previous studies (Boschma et al., 2015; Rigby, 2015; Balland et al., 2018) and build up a *Relatedness Density* indicator for technology *c* in region *r*, RD_{rct}, by using the relatedness function of the EconGeo R package (Balland, 2017).

In analytical terms, having defined the (cognitive) relatedness coefficients between the focal technology *c* and each other technology *i*, φ_{ci}, in terms of co-occurrence, RD_{rc} is defined as their (normalised) weighted sum, using as weights the binary indicators of region *r* 's Revealed Technological Advantages in technologies *i*, as from Eq.(1) (RTA_{ri}):

[6] Results, available from the authors upon request, do not change when a different lag is used.

$$RD_{rc} = \frac{\sum_{i \neq c} RTA_{ri} \cdot \varphi_{ci}}{\sum_{i \neq c} \varphi_{ci}} \cdot 100 \qquad (3)$$

The set of explanatory variables of the regional capacity to enter new technologies is completed by controlling for: the total number of technological claims in the region ($Tech\ Stock_{rt-1}$), as a proxy of regional technological size; the total number of technological claims for each IPC4 code ($Tech\ Size_{ct-1}$), as an indicator of its weight in the knowledge space; regional total employment ($Employment_{rt-1}$) as a proxy of regional economic size (Balland et al., 2018).

Table 1 reports the main descriptive statistics of the variables we have defined.

Table 1 – Descriptive statistics

Variable	Mean	s.d.	Min	Max
New RTA	0.085	0.278	0	1
Relatedness Density (RD)	25.960	7.224	0	100.00
SPI	0.326	0.380	0	1
No. of forward citations	1.08	6.52	0	601.5
TechStock	1.410	1.948	0	12.637
TechSize	1.252	2.389	0	27.206
Employment	4.242	3.414	0.060	24.436

Number of observations is 340,960. TechStock and TechSize are measured in no. of patents. Employment is measured in no. of employees.

3.3 Regression models

Given the dichotomic nature of our dependent variable, we account for its determinants by estimating a Linear Probability Model for the probability that region *r* develops a new RTA in a given technology *c*.[7] In order to test our Hp1 (Section 2), we start by estimating a baseline model (Eq.(4)), in which we augment the technological branching framework. Accordingly, we consider the technological relatedness of the new technology, *c*, RD_{rct-1} as the main regressor, and we add to it our focal explanatory variable, SPI_{rct-1}, and our set of control variables, in vector X_{rct-1}:

$$NewRTA_{rct} = \alpha + \beta_0 RD_{rct-1} + \beta_1 SPI_{rct-1} + \beta_2 X_{rct-1} + Y_t + R_r + C_c + \varepsilon_{rct} \qquad (4)$$

[7] As a robustness check, we also estimate our model using a Logit estimator obtaining very similar results (see Table 4 in Section 4.1).

In Eq.(4), Y_t, R_r, and C_c are year-, region-, and technology-specific fixed effects that account for, respectively, common annual and technology shocks and regional time-invariant unobserved characteristics. ε_{rtc} is an idiosyncratic error term.

In order to test our Hp2, we further augment the baseline model (Eq.(4)) by plugging in it the interaction term between RD_{rct} and SPI_{rct-1}, as in the following model:

$$NewRTA_{rct} = \alpha + \beta_0 RD_{rct-1} + \beta_1 SPI_{rct-1} + \beta_2 (RD_{rct-1} \times SPI_{rct-1}) + \beta_3 X_{rct-1} + Y_t + R_r + C_c + \varepsilon_{rct} \quad (5)$$

In all model specifications, we cluster standard errors at the regional level.

Based on our discussion in the previous section, Hp1 is confirmed if, in Eq.(4) and Eq.(5), β_1 is significant an negative. Hp2 is confirmed if relatedness negatively moderates (i.e., attenuates) the relationship between *SPI* and *NewRTA*, that is, if β_2 is significant and negative. As we will see in the following, in order to refine our analysis about the role of SPIs, we will investigate how its marginal effect from Eq.(5) varies by considering different values of RD_{rct}.

4. Results

Table 2 reports the results of our estimates for the baseline model (Column 1), using SPI as a focal regressor, to which we progressively add region (Column 2), year (Column 3) as well as technology-group fixed effects (Column 4).

Before moving to the results about our SPI regressor, let us notice that in all the specifications, including our preferred and more demanding one of Column 4, results confirm the technological branching framework we have adopted. The entry in a focal region of a new technology is facilitated by its relatedness to the technologies that pre-exist in the same region: RD is in fact positive and strongly significant. Confirming the basic insights of geography of innovation, technological diversification is actually of related nature, as new technologies are developed in a place-dependent way. Still consistently with expectations, the acquisition of a new revealed technological advantage appears more likely to occur in regions that can benefit from a larger endowment of technological knowledge in general terms, and by the technological size of the focal technology: both *TechStock* and *TechSize* are in fact significantly positive. As for the other non-technological control, total regional employment negatively correlates with technological diversification in Column 4. This is somehow unexpected and suggests that, while possibly less endowed with tangible and intangible resources and capabilities, smaller regions might be also more prone and flexible in modifying their knowledge base for accommodating the entry of a new technology.

Coming to our SPI indicator, as we did expect in formulating our Hp1, it shows a significant and negative correlation with *NewRTA*. A larger endowment of mono-IPC patents in the regional technologies does correlate with a reduced capacity to specialise in them, conditionally on their relatedness to pre-existing technologies. Consistently with our interpretative framework, an increase in the siloed nature of the regional knowledge base for a certain technology might in fact reduce the scope of recombining existing knowledge for the acquisition of a new revealed comparative advantage in it.

Table 2 – New Revealed Technological Advantages (NewRTA): relatedness (RD) and Single Patent-code Inventions (SPI)

	(1) New RTA	(2) New RTA	(3) New RTA	(4) New RTA
RD	0.005***	0.013***	0.014***	0.014***
	[0.000]	[0.001]	[0.001]	[0.001]
SPI	-0.017***	-0.019***	-0.019***	-0.019***
	[0.001]	[0.001]	[0.001]	[0.001]
TechStock	-0.003***	0.009***	0.010***	0.010***
	[0.001]	[0.002]	[0.002]	[0.002]
TechSize	0.004***	0.003***	0.003***	0.002***
	[0.000]	[0.000]	[0.000]	[0.000]
(log)Employment	-0.004	-0.036***	-0.038**	-0.039**
	[0.002]	[0.013]	[0.017]	[0.017]
Constant	-0.033***	-0.231***	-0.237***	-0.239***
	[0.006]	[0.020]	[0.023]	[0.024]
Observations	340,961	340,960	340,960	340,960
R-squared	0.014	0.027	0.027	0.029
Region FE	NO	YES	YES	YES
Year FE	NO	NO	YES	YES
Tech. class FE	NO	NO	NO	YES

Notes: Coefficients represent marginal effects. Standard errors, in parentheses, are clustered on regions. SPI is the fraction of Single Patent-code Inventions over total patents. Coefficients for employment, TechStock and TechSize are multiplied by 1,000. Significance level: *p<0.10, **p<0.05, ***p<0.01.

A second important result concerns the augmented specification of the model that, in addressing our Hp2, includes the interaction term between technological relatedness (RD) and SPI. In apparent contradiction with our Hp1, the sign of SPI turns out to be significantly positive now, suggesting that a more siloed knowledge base for a technology could even make its entry in the regional knowledge base more possible. This switch in the result with respect to Table 2 is evidently due to the consideration of the different levels of relatedness at which SPI exerts its effect on NewRTA. For an average level of relatedness (i.e., RD), our Hp1 is confirmed. However, when its specific values are considered, through the inclusion of the interaction between RD and SPI, the picture becomes more nuanced. To start with, by confirming our Hp2, the interaction at stake is significantly negative.[8] As we have suggested in Section 2, targeting more related technologies in diversifying, and thus benefiting from similar capabilities to existing ones in doing that, makes the lack of knowledge combinatorial opportunities entailed by higher SPIs less hampering. In brief, an increasingly higher level of relatedness reduces the average impact that SPI has on regional technological diversification.

Table 3 – New Revealed Technological Advantages (NewRTA): Relatedness (RD), Single Patent-code Inventions (SPI), and their interaction

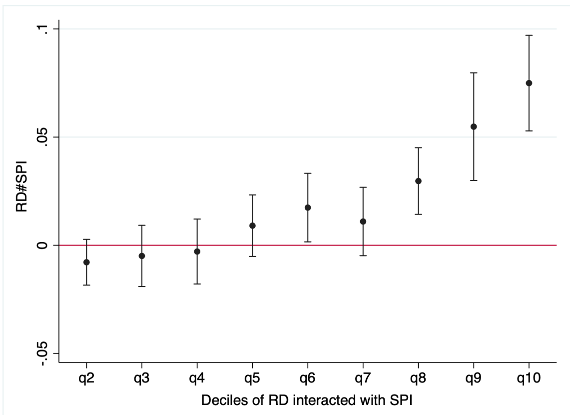
	(1)	(2)	(3)	(4)
VARIABLES	New RTA	New RTA	New RTA	New RTA
RD	0.006***	0.014***	0.014***	0.015***
	[0.000]	[0.001]	[0.001]	[0.001]
SPI	0.030***	0.028***	0.028***	0.027***
	[0.005]	[0.005]	[0.005]	[0.005]
RD#SPI	-0.002***	-0.002***	-0.002***	-0.002***
	[0.000]	[0.000]	[0.000]	[0.000]
TechStock	-0.003**	0.009***	0.010***	0.010***
	[0.001]	[0.002]	[0.002]	[0.002]
TechSize	0.004***	0.003***	0.003***	0.002***
	[0.000]	[0.000]	[0.000]	[0.000]
(log)Employment	-0.004*	-0.037***	-0.040**	-0.041**
	[0.002]	[0.013]	[0.017]	[0.017]
Constant	-0.048***	-0.247***	-0.252***	-0.254***
	[0.005]	[0.021]	[0.023]	[0.024]
Observations	340,961	340,960	340,960	340,960
R-squared	0.014	0.028	0.028	0.030
Region FE	NO	YES	YES	YES
Year FE	NO	NO	YES	YES
Tech. class FE	NO	NO	NO	YES
Diff.:	42.46	35.18	35.35	34.94
p-value	0.0000	0.0000	0.0000	0.0000

Notes: Coefficients represent marginal effects. Standard errors, in parentheses, are clustered on regions. SPI is the fraction of Single Patent-code Inventions over total patents. Coefficients for employment, TechStock and TechSize are multiplied by 1,000. Significance level: *p<0.10, **p<0.05, ***p<0.01.

[8] The interacted coefficient statistically differs from the single estimated coefficient of SPI as confirmed by the Wald test reported in Table 3.

In the same respect, interesting results emerge by observing how the marginal effect of SPI varies for different level of relatedness, as Figure 4 shows with respect to different deciles of the RD distribution (with the first decile as reference category). Quite interestingly, a progressively higher level of relatedness would seem to make the marginal effect of SPI less negative, but this is not significant until very high levels of RD are retained. For the last two deciles of relatedness, a higher incidence of SPI in a regional technology makes its entry in the relative knowledge space more possible, thus accounting for the positive sign of SPI in Table 3. In other words, a siloed knowledge base does not seem to affect the regional diversification in technologies marked by very low or average values of relatedness. This might be so because, with such a (low) level of relatedness, the constraints posed by cognitive proximity in recombining knowledge for the sake of diversification, might be so low to neutralise the lack of combinatorial opportunities of siloed technologies. Conversely, very high levels of relatedness come to ease the entry of a new technology. This is interesting and still possibly consistent with the recombinant theory of innovation at the regional level. When regions target new technologies that draw on very similar capabilities to existing ones, and thus opt for a very narrow case of technological branching, a more siloed knowledge base serves to increase the degree of cognitive homogeneity between to-be-combined knowledge, thus reinforcing the diversification enabling role of relatedness. At (very) high level of relatedness, this even counterbalance the reduction in combinatorial opportunities that SPIs entail and make them facilitate the relative technological diversification.

Figure 4 – New Revealed Technological Advantages (NewRTA):
Single Patent-code Inventions (SPI) for different deciles of relatedness (RD)



Notes: marginal effects with respect to the omitted category (first decile of RD). Confidence intervals at 95%.

4.1 Robustness checks

The results we have obtained are robust with respect to different checks we have implemented. The first robustness check that we propose refers to the econometric estimator we have employed to estimate the relationship between SPI and NewRTA. Since our dependent variable is binary, we have run our model using a Logit estimator. The main results about RD and SPIs, presented in Table 4, are qualitatively identical to those obtained using a Linear Probability Model, confirming our main finding about Hp1. However, the interaction between SPI and RD is not significant now, possibly because of the reduced spectrum of individual RD values for which SPI exerts its effect.

Commentato [CC4]: CI PUò STARE? COME SI PUO' SPIEGARE ALESSANDRO: NON SO SE INTERPRETO BENE LA TUA SPIEGAZIONE, IL FATTO CHE NON E' SIGNIFICATIVO (E POSITIVO) CON LOGIT CI CREA PROBLEMI MA IO QUI STRESSEREI PRINCIPALMENTE IL RISULTATO FORTE SU SPI, ANCHE CONSIDERANDO CHE L'EFFETTO INTERAGITO SI APPREZZA MEGLIO QUANDO SPECIFICATO IN MODO NON LINEARE (GRAFICO), CHE PASSA DA NEGATIVO A POSITIVO. FORSE LOGIT CATTURA IN MODO DIVERSO LA DISTRIBUZIONE DI SPI E RD IN MODO TALE CHE LA MEDIA GENERALE RISULTANTE SIA NON SINGIFICATIVO? HO QUESTA INTUIZIONE MA NON SAPREI SPIEGARLO TECNICAMENTE. FORSE POSSIAMO MOSTRARE IL GRAFICO INTERAGITO ANCHE CON LOGIT SPERANDO CHE IL PATTERN SIA UGUALE?

Table 4 – New Revealed Technological Advantages (NewRTA): Relatedness (RD), Single Patent-code Inventions (SPI) and their interaction

Logit Estimates

	(1) New RTA	(2) New RTA
RD	0.2160***	0.2155***
	[0.003]	[0.003]
SPI1	-0.1451***	-0.1653***
	[0.005]	[0.021]
RD#SPI		0.0006
		[0.001]
TechStock	0.1750***	0.1751***
	[0.013]	[0.013]
TechSize	0.0820***	0.0826***
	[0.002]	[0.002]
(log)Employment	-0.3861***	-0.3844***
	[0.131]	[0.131]
Constant	-7.5462***	-7.5341***
	[0.262]	[0.262]
Observations	338,766	338,766
Region FE	YES	YES
Year FE	YES	YES
Tech. class FE	YES	YES
Diff. p-value		0.000

Logit estimates. Standard errors, in parentheses, are clustered on NUTS-2 regions. Significance level: *p<0.10, **p<0.05, ***p<0.01.

The second robustness check we run is more substantial and refers to the hypotheses, implicitly made so far, that SPIs are characterised by a homogenous knowledge content across them. This represents a simplifying assumption that does not reflect the (actual) highly heterogenous quality distribution of patent inventions (Hall et al., 2005; Lanjouw and Schankerman, 2004, among others). Despite the number of SPIs is historically lower than multi-code patents, if SPIs systematically embodied a superior knowledge content and a greater invention quality, the regional capacity of diversifying (Hp1) and that of taking stock of relatedness in entering new technologies (Hp2) may be differently affected by the presence of SPIs. To test the validity of our findings, we thus account for heterogenous SPI quality distribution using forward patent citations as a proxy for patent quality. The number of citations received by each patent in fact represents the most notable standard indicator to measure the success and quality of the innovative content of patents (Criscuolo and Verspagen, 2008; Hall et al., 2005; Squicciarini et al., 2013). Operationally, we have collected patent-level data on forward citations received at five years after publication, included in the Patent Quality Indicators Database provided by the OECD (Squicciarini et al., 2013), and we have re-run our estimates by using citation-weighted SPIs.

Table 5 shows the results of this further model specification. Consistently with our previous estimates, when all the FEs are retained (Column 1), SPI reveals a negative significant correlation with NewRTA for an average level of RD (though only weakly significant). Drawing on the discussion we have made in Section 2 on this point, it seems that the cognitive linkages that SPIs establish in terms of citations, because of their inner quality, and the effect these linkages could have in the generation of knowledge recombination opportunities, are not sufficient to compensate the hampering role that SPIs have in regional technological diversification. However, still consistently with the baseline model, the interaction between citation-weighted SPI and RD is again significantly negative in Column (2) and, once more, by making RD varies from its average value, the sign of (weighted) SPI switches from negative to positive.

Table 5 – New Revealed Technological Advantages (NewRTA): Relatedness (RD), Single Patent-code Inventions (SPI) and their interaction weighted by forward citations

	(1) New RTA	(2) New RTA
RD	0.001*	0.002***
	[0.000]	[0.001]
SPI (weighted)	-0.025***	0.035***
	[0.004]	[0.011]
RD#SPI (weighted)		-0.002***
		[0.000]
TechStock	0.000	0.001
	[0.001]	[0.001]

TechSize	-0.001**	-0.001**
	[0.000]	[0.000]
(log)Employment	0.018	0.017
	[0.015]	[0.015]
Constant	0.014	-0.024
	[0.025]	[0.027]
Observations	340,960	340,960
R-squared	0.011	0.012
Region FE	YES	YES
Year FE	YES	YES
Tech. class FE	YES	YES
Diff.:		10.4324
p-value		0.0014

Estimates are weighted by the number of citations received in 5 year time after publication. Coefficients represent marginal effects. Standard errors are clustered on regions. Coefficients for employment TechStock and TechSize are multiplied by 1000. Significance level: *p<0.10, **p<0.05, ***p<0.01.

5. Conclusions

Diversification is a crucial leverage for regions to face the constant pressure of renew their technological specialisations, which exogenous and endogenous shocks are posing to them. Its analysis has been conspicuous in the last decades, showing that regions normally diversify, and should diversify, their set of technologies in a related way, by taking stock of pre-existing knowledge and searching for new recombination of it. In the same stream of research, several factors have been however identified that intervene in the unfolding of technological branching and make it occur with several nuances (see Boschma, 2016, for a review). To the best of our knowledge, this kind of "augmented" analysis of the relatedness approach has so far neglected the role of the siloed nature of the knowledge base that underlies regional technologies, as it is revealed by the diffusion and incidence of SPIs in local technologies.

In this paper we fill in this gap, starting from the empirical recognition that the diffusion of Single-Patent-code-Inventions, as a proxy of siloed knowledge base technologies, is far from exiguous. We then argue that this aspect of the geography of innovation could affect the way in which technological diversification unfolds at the regional level and show that this is so in an empirical application to EU28 NUTS2 regions over the period 1986-2017.

As we did expect from a theoretical point of view, the siloed knowledge of technologies appears to pose regions an important challenge. With respect to a reduced spectrum of multi-domain inventions per technology, the existing knowledge base of the region becomes less rich of knowledge-recombination opportunities and this reduces the scope of technological diversification.

However, as we also expected, this effect is less in place with respect to technologies that are cognitively closer to the extant regional ones. To be sure, the SPI effect is significant only for high levels of such a relatedness, for which it switches from negative into positive. As we said, this is interesting and still possibly consistent with the recombinant theory of innovation at the regional level, pointing to the fact that, conditionally on the level of relatedness, the effect of a siloed knowledge base on diversification is in fact twofold. Not only a more siloed knowledge base reduces the scope of knowledge recombinations leading to diversification in a focal technology, but it also makes the knowledge to be combined more homogeneous and, for very related technologies, this helps in diversifying.

These results suggest that an additional aspect emerges in the implementation of Smart Specialisation Strategies (S3), which are based on and recommend to regions a related process of technological diversification. According to the main S3 rationale (Boschma and Giannelle, 2014), targeting new technologies that are cognitively closer to pre-existing ones enable regions to escape the risks and costs of entering unfamiliar technological domains, of which they have little or no capabilities. To this standard S3 rationale, our study adds a further new one. Sticking to relatedness in diversifying can also be a way to neutralise the drawbacks of pursuing technologies that have a siloed knowledge base and thus lack of knowledge combinatorial opportunities. Even more, it serves to turn the "simple" nature of these technologies into an advantage for the sake of diversification.

This is an important argument, which regional policy makers will have to retain, especially in those local contexts that pervasively host technologies marked by a siloed knowledge base. Indeed, this is the most important research avenue that our findings stimulate. Future research should try to investigate which are the regions that are more prone to host SPIs in developing their technologies, starting from the standard distinction between core vs. peripheral regions. At a more methodological level, other proxies of the siloed knowledge base of local technologies than the incidence of SPIs should be investigated, possibly by considering the extent to which citations, in addition to inventions, are marked by single patent code patterns. As usual, the most delicate issue however remains that of the possible endogeneity of our focal regressor, which we have only superficially considered by lagging it with respect to the dependent variable and by using a rich set of fixed effects in our econometric model. Still, future research should look for a more exogenous treatment of SPIs, which could lead us to consider its correlation with technological branching as an actual causal relationship.

References

Acs, Z. J., Anselin, L., and Varga, A. (2002). Patents and innovation counts as measures of regional production of new knowledge. Research Policy 31 (7): 1069–85. doi:10.1016/S0048-7333(01)00184-6.

Asheim, B.T. and Coenen, L. (2005). Knowledge bases and regional innovation systems: Comparing Nordic clusters. Research Policy, 34: 1173-1190.

Asheim, B. (2007), Differentiated knowledge bases and varieties of regional innovation systems, Innovation: The European Journal of Social Science Research, Volume 20(3).

Balland, P.A. (2016). Chapter 6: Relatedness and the geography of innovation, in Shearmu, R., Carrincazeaux, C. and Doloreux. D. (eds), Handbook on the Geographies of Innovation. London, Edward Elgar.

Balland, P. A. (2017). EconGeo: Computing key indicators of the spatial distribution of economic activities, R package version 1.3. Retrieved from https://github.com/PABalland/EconGeo

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, DOI: 10.1080/00343404.2018.1437900.

Balland, P. A. and Rigby, D. (2017). The geography of complex knowledge. Economic Geography, 93(1), 1–23. doi:10.1080/00130095.2016.1205947.

Becker, M.C., Knudsen, T. and Swedberg, R.R. (2012). Schumpeter's theory of economic development: 100 years of development, Journal of Evolutionary Economics 22, 917–933. doi:10.1007/s00191-012-0297-x.

Boschma, R.A. & Frenken K. (2006). Why is economic geography not an evolutionary science? Towards an evolutionary economic geography, Journal of Economic Geography, 6 (3): 273–302.

Boschma, R., Minondo, A. & Navarro, M. (2013). The emergence of new industries at the regional level in Spain: a proximity approach based on product relatedness, Economic Geography, 89, 29–51.

Boschma, R., and Giannelle, C. (2014). Regional branching and smart specialisation policy. S3 Policy Brief Series 06/2014. Seville: European Commission.

Boschma, R., Balland, P. A., & Kogler, D. F. (2015). Relatedness and technological change in cities: the rise and fall of technological knowledge in US metropolitan areas from 1981 to 2010. Industrial and corporate change, 24(1), 223-250.

Breschi, S., Malerba, F. and Orsenigo, L. (2000). Technological regimes and Schumpeterian patterns of innovation. The Economic Journal, Volume 110, Issue 463, April 2000, Pages 388–410, https://doi.org/10.1111/1468-0297.00530.

Cicerone, G., Faggian, A., Montresor, S., Rentocchini, F. (2019). Harnessing artificial intelligence for regional eco-innovation: a patent-based analysis of European regions' green-tech diversification. Paper prepared for the 2019 AISRE Conference.

Content, J., & Frenken, K. (2016). Related variety and economic development: a literature review. European Planning Studies, 24(12), 2097-2112.

Criscuolo P.; Verspagen B. (2008). Does it matter where patent citations come from? Inventor vs. examiner citations in European patents. Research policy, 37(10), 1892-1908.

Dosi, G. (1982). Technological paradigms and technological trajectories: A suggested interpretation of the determinant and direction of technological change, Research Policy, 11, 147-162.

Essletzbichler, J. (2015) Relatedness, Industrial Branching and Technological Cohesion in US Metropolitan Areas, Regional Studies, 49:5, 752-766,

Evenson, R. and Kislev, Y. (1976). A stochastic model of applied research, Journal of Political Economy 84, 265–282.

Fleming, L., Sorenson, O. (2001). Technology as a complex adaptive system: evidence from patent data. Research Policy 30 (7), 1019–1039.

Farhauer, O. and Kröll, A. (2012). Diversified specialisation—going one step beyond regional economics' specialisation-diversification concept, Jahrb Reg wiss, 32:63–84

Frenken, K., Van Oort, F., & Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. Regional studies, 41(5), 685-697.

Jacobs, J. (1969). The Economy of Cities, New York: Vintage Books

Janssen, M. J., & Frenken, K. (2019). Cross-specialisation policy: rationales and options for linking unrelated industries. Cambridge Journal of Regions, Economy and Society, 12(2), 195-212.

Kogut, B. and Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. Organization Science, 3, 383–397. doi:10.1287/orsc.3.3.383.

Hall, B. H., Jaffe, A., Trajtenberg, M. (2005). Market value and patent citations. RAND Journal of economics, 16-38.

Hassink, R. and Gong, H. (2019). Six critical questions about smart specialization, European Planning Studies, DOI: 10.1080/09654313.2019.1650898.

Hausmann, R. & Klinger, B. (2007). The structure of the product space and the evolution of comparative advantage, CID Working Paper, 146. Cambridge, MA: Center for International Development, Harvard University.

Hidalgo, C. and Hausmann, R. (2009). The building blocks of economic complexity. Proceedings of the National Academy of Sciences, 106, 10570–10575. doi:10.1073/pnas.0900943106.

Hidalgo, C., Balland, P.A., Boschma, R., Delgado, M., Feldman, M., Frenken, K., Glaeser, E., He, C., Kogler, D., Morrison, A., Neffke, F., Rigby, D., Stern, S., Zheng, S., & Zhu, S. (2018). The Principle of Relatedness, Springer Proceedings in Complexity, 451-457.

Lanjouw, J. and M. Schankerman (2004). Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators. The Economic Journal, 114(495): 441-465.

Malerba, F. and Orsenigo, L. (1996). Schumpeterian patterns of innovation are technology-specific, Research Policy, 25, 451-478.

Castellani, D., Marin, G., Montresor, S., & Zanfei, A. (2022). Greenfield foreign direct investments and regional environmental technologies. Research Policy, 51(1), 104405. Mewes, L. (2019). Scaling of atypical knowledge combinations in American metropolitan areas from 1836 to 2010. Economic Geography, DOI: 10.1080/00130095.2019.1567261.

Montresor, S. and Quatraro, F. (2017). Regional branching and Key Enabling Technologies: Evidence from European Patent data, Economic Geography, 93:4, 367-396, DOI: 10.1080/00130095.2017.1326810.

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.

Neffke, F. & Henning, M. (2013). Skill relatedness and firm diversification. Strategic Management Journal, 34(3), 297-316.

Nelson, R.R. and Winter, S. (1982). An Evolutionary Theory of Economic Change, Cambridge, Massachusetts, The Belknap Press of Harvard University Press.

Rigby, D. (2015). Technological relatedness and knowledge space: Entry and exit of US cities from patent classes. Regional Studies,49(11), 1922–1937. doi:10.1080/00343404.2013.854878.

Rosenberg, N. (1976). Perspectives on Technology, Cambridge: Cambridge University Press.

Santoalha, A. (2019). Technological diversification and Smart Specialisation: the role of cooperation, Regional Studies, 53:9, 1269-1283, DOI: 10.1080/00343404.2018.1530753.

Schmoch, U. (2008). Concept of a technology classification for country comparisons. Final report to the world intellectual property organisation (wipo), WIPO.

Squicciarini, M., Dernis, H. and Criscuolo, C. (2013). Measuring Patent Quality: Indicators of Technological and Economic Value. OECD Science, Technology and Industry Working Papers, 2013/03, OECD Publishing, Paris.

Tanner, A. N. 2014. Regional branching reconsidered: Emergence of the fuel cell industry in European regions. Economic Geography 90 (4): 403–27. doi:10.1111/ecge.2014.90.issue-4.

Verhoeven, D., Bakker, J. and Veugelers, R. (2016), Measuring technological novelty with patent-based indicators, Research Policy, 45(3), 707-723.

Weitzman, M. L. (1998). Recombinant growth. Quarterly Journal of Economics 113 (2): 331–60. doi:10.1162/003355398555595.

Winter, S.G. (1984). Schumpeterian competition in alternative technological regimes, Journal of Economic Behaviour and Organization, 5, 287-320.

Winter, S.G. (1987). Knowledge and competence as strategic assets, in Teece D. (ed.), The Competitive Challenge. Strategies for Industrial Innovation and Renewal, Cambridge, Mass., Ballinger Publishing Company.

Xiao, J., Boschma, R. & Andersson, M. (2018a). Industrial diversification in Europe: The differentiated role of relatedness, Economic Geography, 94:5, 514-549. DOI:10.1080/00130095.2018.1444989.

Zhu, S., Hey, C. & Zhou, Y. (2017). How to jump further and catch up? Path-breaking in an uneven industry space, Journal of Economic Geography, 17, 521–545

Appendix

Table A1 – Correlation matrix

	New RTA	RD	SPI	RD#SPI	Tech. Stock	Tech. Size	Empl. (log)
New RTA	1						
RD	0,11	1					
SPI	-0,01	0,07	1				
RD#SPI	0,09	0,29	0,93	1			
Tech. Stock	0,02	0,39	0,04	0,14	1		
Tech. Size	0,01	-0,13	0,08	0,04	-0,07	1	
Empl. (log)	0,03	0,48	0,05	0,16	0,59	-0,07	1