THE ART OF CHOOSING BATTLES: THE NEED FOR SELECTIVITY IN SMART SPECIALIZATION STRATEGIES

By Zenne Hellinga
Supervisor: Dr. Pierre-Alexandre Balland
External supervisor: Julia Bachtrögler-Unger, PhD
April 2023

Abstract

More than €62 billion in funding for Research & Innovation was allocated via the Smart Specialization Strategy. Therewith, smart specialization may well be one of the biggest multinational strategies aiming to boost innovation ever. Various scholars have pointed out that regional S3 strategies target too many broad areas of intervention. This paper shows how this lack of selectivity undermines the correspondence between S3 priorities and regional technological capabilities, the alignment between S3 priorities and R&I funding, and the relation between R&I funding and knowledge production. The results indicate that concentrating R&I funding on a limited set of technological domains increases its effectiveness in terms of knowledge production.

Utrecht University

1. INTRODUCTION

The Smart Specialization Strategy (S3) plays a vital role in the EU's Cohesion Policy. More than €62 billion in funding for Research & Innovation was allocated via the S3 policy in Europe between 2014-2020 (European Commission, 2022). Therewith, smart specialisation may well be one of the biggest multinational strategies aiming to boost innovation ever (Asheim et al., 2017).

Its relatively short journey from concept to policy has given it the reputation of being "a perfect example of a policy running ahead of theory" (Foray et al., 2011, p. 1). Thus far, only a few scholars begun examining the policy's first programming period (2014-2020) by assessing the correspondence between regional implementations of the S3 and each region's intrinsic characteristics. Several point out that regional S3 strategies lack selectivity in their targeted areas of intervention (Iacobucci & Guzzini, 2016; Giannelle et al., 2020a; Di Cataldo et al., 2021; Kramer et al., 2021; Marrocu et al., 2023), although the policy prescribes that resources should be concentrated on a limited set of research and innovation priorities (Foray et al., 2012).

This paper shows how this lack of selectivity undermines S3's implementation and effectiveness by examining three aspects of the policy. 1) The correspondence between regional S3 priorities and regional technological capabilities. 2) The correspondence between regional S3 priorities and regional R&I funding through the European Regional Development Fund (ERDF). 3) The relation between ERDF R&I funding and knowledge production. I do so through a novel integration of three large datasets (Eye@RIS3, ERDF project data, and OECD's REGPAT dataset). Therewith, I will answer the following research questions:

1. To what extent do a region's prioritized technological domains reflect regional capabilities?
2. To what extent do regions spend ERDF R&I funding in line with prioritized technological domains?
3. Is there a relation between ERDF R&I funding and knowledge production in the associated technological domains?

The remaining part of this paper is structured as follows. The next section discusses the concept of S3, the scientific debate surrounding it, and the evaluations of its implementation conducted thus far. The third section describes the different sources of data used in this study and how these were integrated. Section 4 examines the results of the analysis of the three research questions. The last section concludes by discussing this study's limitations, its implications, and directions for future research.

2. THEORETICAL BACKGROUND

2.1 WHAT IS S3?

The concept of smart specialization stems from the idea that the evolution of a regional innovation system is inherently linked to its context. A region's development path depends on its ongoing economic dynamics and institutional structures. This is aptly illustrated by Boschma & Wenting (2007) in their study on the spatial evolution of the British automobile industry. They show that the automobile industry mostly thrived in the regions that were already heavily involved in industries closely related to the car industry, such as bicycle or coach making. Hence, regions should build on their existing strengths and capabilities.

Therefore, S3 diverges from a top-down one-size-fits-all policy towards a bottom-up innovation policy that is tailor-made to each region. Every region should concentrate its public resources on a limited set of well-defined economic, technological or scientific domains, in which it either shows a competitive advantage or a considerable growth potential (Foray et al., 2012).

These targeted domains are called *priorities* or *priority areas* and are identified via the *Entrepreneurial Discovery Process* (EDP) (Foray et al., 2012). The involvement of local entrepreneurial actors in the process of discovering priority areas is a key feature of S3. Its underlying rationale can be traced back to Storper's (1997) idea of regional economies as stocks of relational assets, i.e., local communities with their very own conventions, practices and (tacit) knowledge. Given their direct involvement in such communities, local entrepreneurial actors are probably better equipped than policymakers to understand these communities and identify the most promising paths for regional diversification (Foray et al., 2009; D'Adda et al., 2019).

A systematic identification of these diversification paths is a key challenge for smart specialization. Balland et al. (2019) developed a theoretical framework for regions to identify the most promising areas for smart specialization. Central to this framework are the concepts of relatedness and knowledge complexity. It is most interesting for regions to diversify into highly complex technologies, since these are hard to imitate, therefore sticky in space (Balland et al., 2020), and so, are expected to generate the highest long-term profits. Unfortunately, the knowledge and capabilities needed to diversify into these complex technologies, are hard to attain. Therefore, regions have the most potential to develop new complex technologies in those areas related to existing capabilities (Hidalgo et al., 2018), as shown in *Figure 1*. There is a large strand of literature showing empirical evidence in support of this framework (Hausmann et al., 2006; Hidalgo & Hausmann, 2009; Neffke et al., 2011; Essletzbichler, 2013; Kogler et al., 2013; Rigby, 2013; Boschma et al., 2015; Petralia et al. 2022; Rigby et al., 2022; Antonelli et al., 2022; Mewes & Broekel, 2022).

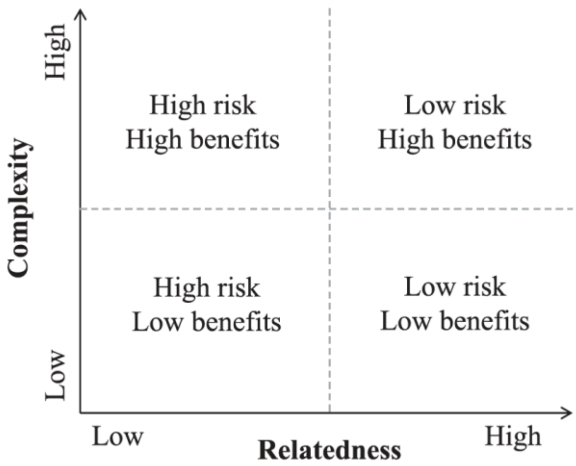

Figure 1. Framework for smart specialization (Balland et al., 2019).

2.2 Concerns and critiques about S3

Smart specialization gained its importance in a relatively short period of time. It was first coined by the Knowledge for Growth Expert Group around 2009 (Foray, 2014) and initially implemented only 5 years later. This short journey from concept to policy raised a range of concerns among experts.

Hassink & Gong (2019) argue that the smart specialization became to be an umbrella term for various concepts in economic geography. Therefore, it is often not well understood by those responsible for its implementation (Kroll, 2015; Capello & Kroll, 2016; Pugh, 2018). Others are concerned that involving local stakeholders in regional innovation policy entails the risk of rent-seeking behaviour, corruption, and regional lock-ins (Camagni & Capello, 2013; Boschma, 2014; Rodríguez-Pose et al., 2014; Grillitsch, 2016; Trippl et al., 2020).

Moreover, there are several concerns about S3 related to governance institutions. Rodríguez-Pose et al. (2014) show that a high quality of regional governmental institutions seems to be crucial for a successful implementation. On top of that, the EU is characterized by diverse structures of governance within its Member States. Embedding S3 in these various institutional contexts can be challenging (Kroll, 2015; Capello & Kroll, 2016; Pugh, 2018).

This also relates with Hassink & Gong's (2019) concern of how well S3 works next to already existing innovation policies.

Lastly, several scholars point out that, while the Cohesion Policy tends to strengthen weaker regions, the logic behind S3 favours more advanced regions. The elements crucial for the policy's implementation are exactly those that are missing in lagging regions (Boschma et al., 2014; McCann & Ortega-Argilés, 2015; Capello & Kroll, 2016; Iacobucci & Guzzini, 2016). This may drive these regions to set many broad priorities instead of a few well-defined ones (Boschma, 2014; Capello & Kroll, 2016; Di Cataldo et al., 2021).

2.3 Evaluations of S3 implementation

There is a growing body of literature evaluating the implementation of S3. This section provides an overview and emphasizes where the current paper differentiates. These studies can be divided into two groups: studies examining regional S3 priorities (see *Table 1*), and studies investigating regional S3 strategies and regional funding decisions (see *Table 2*).

Table 1. Overview of studies evaluating S3 prioritization in alphabetic order.

Study	Focus / aim	Geographical scope	Type of priority	Data
Biagi et al. (2021)	The rationale of regions for prioritizing tourism in S3	191 EU regions	Tourism	Regional tourism statistics
Buyukyazici (2022)	Do priorities reflect regional workplace knowledge and skills?	20 Italian regions	Economic domains	Italian Sample Survey on Professions and Italian Labour Force Survey
D'Adda et al. (2019)	Do priorities reflect regional innovative capabilities?	23 Italian regions	Technological domains	Patent data
D'Adda et al. (2020)	The relatedness between priorities	19 Italian regions	Technological domains	Patent data
Deegan et al. (2021)	The relatedness and complexity of priorities	128 European regions	Economic domains	SBS employment data
Di Cataldo et al. (2021)	Do priorities reflect economic characteristics?	All regions and countries in the Eye@RIS3 dataset	Economic and scientific domains, and policy objectives	GDP, population, unemployment, EU QoG index, patents per inhabitant, tertiary educated
Farinha et al. (2020)	Do priorities reflect regional stakeholders' perceptions?	7 Portuguese regions	Priorities in general	Survey among stakeholders
Gianelle et al. (2020a)	How are priorities indicated and described?	39 Italian and Polish regions	Priorities in general	Descriptive analysis of RIS3 documents
Iacobucci & Guzzini (2016)	The relatedness of priorities and their potential interregional links	16 Italian regions	Technological domains	Descriptive analysis of RIS3 documents
Kramer et al. (2021)	Do priorities reflect regional capabilities?	185 European regions	Technological, economic, and scientific domains	Patent data, SBS employment data, and scientific publication data.
Lopes et al. (2018)	Do priorities reflect regional stakeholders' perceptions?	7 Portuguese regions	Priorities in general	Survey among stakeholders
Marrocu et al. (2023)	Do priorities reflect regional capabilities?	243 European regions	Economic domains	SBS employment data
Sörvik & Kleibrink (2015)	What are the most common (combinations) priorities? Do priorities reflect economic characteristics?	174 Eu regions, 18 non-EU regions	Economic domains	Descriptive statistics of Eye@RIS3 data and SBS employment data

As can be seen in *Table 1*, most studies in the first group examine how well regional S3 priorities reflect regional capabilities. They vary in their geographical scope and type of priority, the latter determining the data used (e.g., employment data for economic domains and patent data for technological domains). Moreover, the second column shows that researchers use different indicators to evaluate S3 priorities. Some look at the relatedness between priorities (Iacobucci & Guzzini, 2016; D'Adda et al., 2020), whereas others measure the relatedness of chosen priorities to a region's technological or economic profile (Deegan et al., 2021; Kramer et al., 2021; Marrocu et al., 2023). Location quotients or revealed comparative advantages are used to compare priorities to regional capabilities (D'Adda et al., 2019; Kramer et al., 2021; Marrocu et al., 2023). Following Balland et al.'s (2019) framework, two studies also considered the complexity of prioritized domains (Deegan et al., 2021; Kramer et al., 2021).

Several scholars point out that a systematic analysis of S3 priorities is challenging because priorities are defined in a non-codified way (Iacobucci & Guzzini, 2016; D'Adda et al., 2019; Marrocu et al., 2023). This methodological problem is less of an issue for analysing economic domains since regions also have to select the associated NACE sectors for each priority (see *Data and Methods* section for a more detailed description of S3 priority data). However, this is not the case for the analysis of technological domains. In their analysis of Italian regions, D'Adda et al. (2019) and D'Adda et al. (2020) overcome this issue by matching priorities to three-digit IPC codes using WIPO's automatic categorization assistant[1]. Alternatively, Kramer et al. (2021) use a less fine-grained codifying method, by matching priorities with Schmoch's (2008) technology classes via automatic text mining. By thoroughly analysing all S3 priorities, the current study created a technological taxonomy that is tailored to S3 priority data (see *Data and Methods*). Therewith, the correspondence between prioritized technological domains and a region's technological capabilities can be measured more accurately.

Some syntheses about S3 priorities can be drawn from the studies in *Table 1*. First, Marrocu et al. (2023) show that most regions tend to prioritize economic domains that are not very related to their economic profile. The results of Deegan et al. (2021) indicate a stronger relation between economic relatedness and selected priorities. Second, Deegan et al. (2021) also show that regions tend to prioritize more complex economic domains. However, overall, regions do

[1] Categorization Assistant in the International Patent Classification (IPCCAT), see www.wipo.int/ipccat/.

not consider relatedness and complexity in tandem when selecting priorities, as proposed by Balland et al. (2019), but rather independently. Third, both Deegan et al. (2021) and Di Cataldo et al. (2021) notice that regions often mimic neighbouring regions in their S3 strategies. Fourth, there is a moderate degree of coherence between each region's prioritized technological domains and the domains in which they possess a comparative advantage, with a slightly higher degree for more developed regions (D'Adda et al., 2019; Kramer et al. 2021). This degree tends to be lower for prioritized economic domains (Kramer et al. 2021; Marrocu et al. 2023). Moreover, Kramer et al. (2021) note that the regions that define more broad and vague priorities often have a better correlation between priorities and regional capabilities. This relates to the commonly shared notion that regions are not selective in their priority setting, and that most priority areas are very broadly defined (Iacobucci & Guzzini, 2016; Gianelle et al., 2020a; Di Cataldo et al., 2021; Marrocu et al., 2023). This seems to be more of an issue for regions with a weaker quality of governance (Di Cataldo et al., 2021).

As shown in *Table* 2, a relatively small body of literature is concerned with how S3 influences regional funding decisions. D'Adda et al. (2021) investigate to what extent S3 changed the allocation of structural funds in Italy by comparing S3's first programming period of 2014-2020 with the preceding one. They find that the changes are modest and differentiate among regions. The other studies in this group focus on the alignment between S3 priorities and a region's funding decisions by analysing regional project selection procedures, commonly referred to as *calls for proposals*. In the same vein as D'Adda et al. (2021), both Giannelle et al. (2020a) and Fratesi et al. (2021) conclude that S3 did not engender much change from former more horizontal industry intervention policies. They argue that this is caused by the lack of selectivity in S3 priorities, as well as the fact that most calls for proposals address all a region's S3 priorities collectively. Kramer et al. (2021) add that this is more prevalent in less developed regions.

Table 2. Overview of studies evaluating S3 funding in alphabetic order.

Study	Focus / aim	Geographical scope	Data
D'Adda et al. (2021)	To what extent is the allocation of structural funds changed because of S3?	15 Italian regions	ESF project data, ERDF projects are excluded.
Fratesi et al. (2021)	Are calls for proposals aligned with priorities? Do they favour collaborative projects? Do they stimulate entry into new activities? Do they support stakeholder communities?	6 EU countries and 17 EU regions	Calls for proposals and RIS3 documents
Gianelle et al. (2020a)	How are priorities defined? Are calls for proposals aligned with priorities? How specific are calls for proposals?	21 Italian and 16 Polish regions	Calls for proposals and RIS3 documents
Kramer et al. (2021)	Are calls for proposals aligned with priorities? Are funded projects aligned with priorities?	185 European regions	Calls for proposals, ERDF project data

Additionally, Kramer et al. (2021) study the alignment of S3 priorities and funding decisions by examining the actual projects that are funded by the ERDF. They clearly show that the use of funded projects gives a more accurate picture of how well funding is aligned with S3 priorities. Where they find that 84% of the calls for proposals correspond to S3 priorities, only 54% of the funded projects are aligned. The current piece of work will do something similar by building upon the work of Bachtrögler et al. (2021). However, as argued before, by using a technological taxonomy tailored to S3 priority data, I will be able to provide a more precise view of the alignment between S3 priorities and funded projects.

2.3 The impact of S3

Evidence on the impact of S3 is very limited, mainly because it is not yet possible to measure the policy's structural effects on innovation and regional economies. Rigby et al. (2022) overcome this issue, not by examining the actual impact of S3, but by mapping to what extent EU cities have followed a smart specialization trajectory since 1980. They do this by tracking technological complexity and relatedness of cities, as proposed by Balland et al. (2019). Their results indicate that cities following such a smart specialization trajectory enjoy an economic performance premium over cities that do not.

Additionally, some scholars investigate S3's short term impact. For instance, Romão (2020) studies the effect of S3 on the tourism sector, and find positive effects on tourism demand, supply and specialization in regions that prioritized tourism. Santos et al. (2022) study the effect of S3 on regional productivity in Portugal by comparing the S3's first programming period of 2014-2020 with the preceding one. They find that the policy generates an additional effect on regional productivity, but only when it is combined with other types of innovation subsidies. Likewise, Crescenzi et al. (2020) analyse the effect of an S3 forerunner programme in Italy on firm performance in Italy. Their analysis indicates that the programme was unsuccessful in boosting investments, value added and employment, except for low-tech sectors. However, as the authors conclude, their results are not applicable to other EU regions. An EU-wide analysis of S3's impact is still missing. Moreover, none of the aforementioned literature investigates the policy's impact on technological innovation, although this is one of its main targets.

2.4 HYPOTHESES

Collectively, the studies presented above provide a detailed picture of S3's first programming period. I will move along the same lines, aiming to fill the following gaps. 1) An EU-wide analysis of the consistency of regional S3 priorities with each region's technological capabilities, using a taxonomy that accurately fits S3 priority data. 2) The alignment between R&I-projects funded by the ERDF and prioritized technological domains. 3) The short-term impact of ERDF R&I funding on knowledge production, and how this impact is influenced by the lack of selectivity in regional S3 strategies.

Building upon the body of literature discussed above, I formulate three hypotheses:

Hypothesis 1: A region's prioritized technological domains reflect its technological capabilities.

Hypothesis 2: A region's ERDF R&I expenditure is in line with its prioritized technological domains.

Hypothesis 3: There is a positive relation between a region's ERDF R&I expenditure and knowledge production. The more a region concentrates its expenditure on a particular technological domain, the stronger this relationship becomes.

3. DATA AND METHODS

3.1 TECHNOLOGICAL TAXONOMY

To examine the research questions, a novel integration of the following three datasets is made: Eye@RIS3 dataset (European Commission, 2018), ERDF project data (Bachtrögler et al., 2021), and OECD REGPAT database (OECD, August 2022). This integration is achieved by making use of the taxonomy of technological domains shown in *Figure 2*. This taxonomy is based on a thorough analysis of all regional S3 priorities in the Eye@RIS3 dataset, and therefore aptly fits this dataset. Most of the technological domains are relatively specific, but since some priorities described very broad technological domains (such as 'ICT' or 'sustainable energy') it was necessary to also include broader domains. *Figure 2* shows how the specific domains are connected to the broader ones.

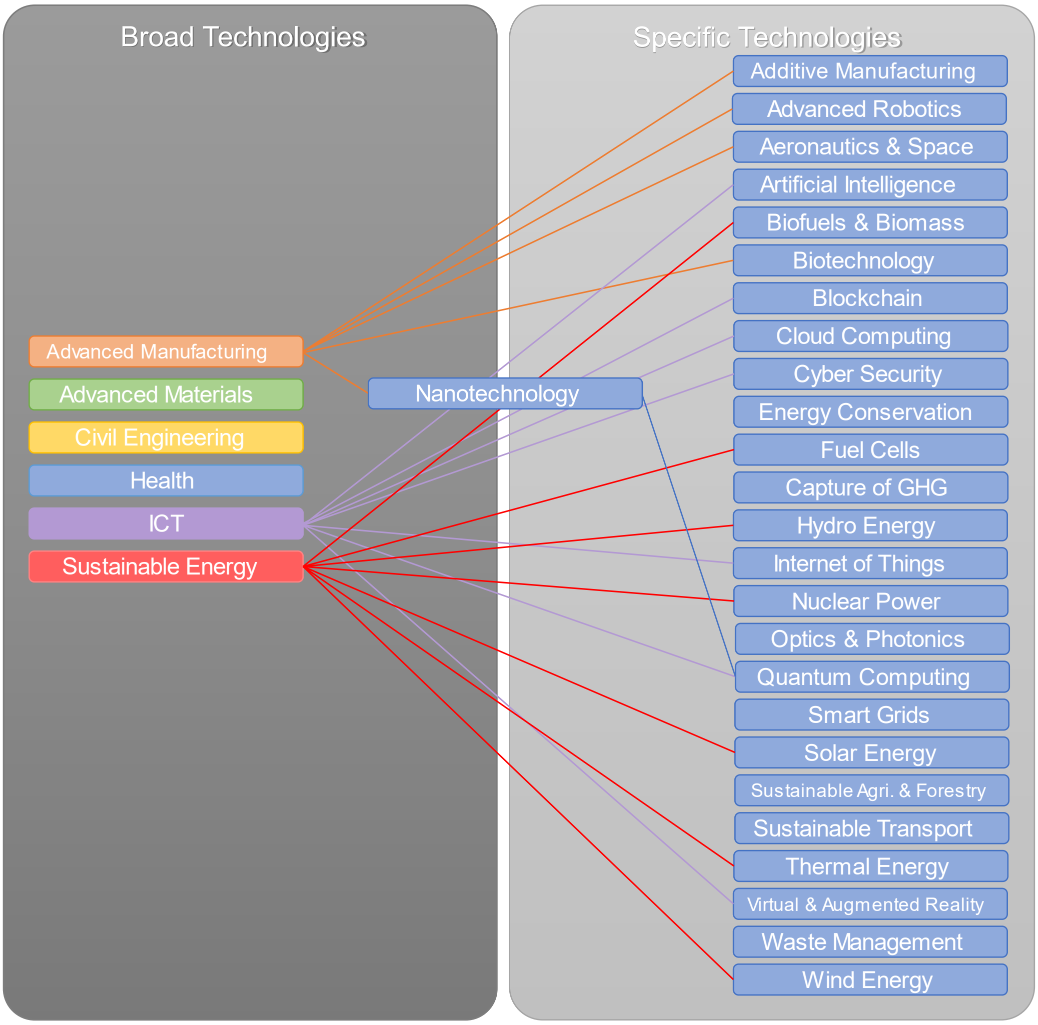

Figure 2. Technological taxonomy

3.2 S3 PRIORITIES (EYE@RIS3 DATASET)

Eye@RIS3 is an online database, available at the European Commission's (EC) S3 Platform[2], containing all the S3 priorities defined by national and regional authorities in RIS3 documents. The strategies differ in territorial level, but most are on a NUTS2 level[3]. Each priority comprises the following pieces of information: the region or Member State, a free text description of the priority, the associated economic domains (based on the 2-digit NACE sectors), the associated scientific domains (based on NABS2007 categories), the associated

[2] https://s3platform.jrc.ec.europa.eu/map

[3] All national strategies were excluded from the analysis and all other strategies were converted to the NUTS2 level to make comparable analysis possible. See *Appendix 7.1* for a description of the used procedure.

policy objectives (based societal grand challenges identified in Horizon2020 and the headline policies in the Innovation Union Flagship Initiative), and the date when a priority was set (see *Appendix 7.1* for a description of how the date and the territorial level of S3 priorities are treated in the econometric analysis).

Priorities were matched to technological domains by combining automatic text mining and manual content analysis of the free text descriptions, and an analysis of the selected economic and scientific domains and policy objectives. Regions identify 5.8 priorities on average, of which 93.7% were covered by the technological taxonomy. As discussed before, regions tend to define very broad priorities. Therefore, each individual priority is associated with 4.6 technological domains on average, which means that the total number of prioritized technological domains averages 15.6. This strongly contradicts the selectivity of intervention areas that S3 advocates, which is a common finding in studies evaluating S3 priorities (Gianelle et al., 2020a; Di Cataldo et al., 2021; Marrocu et al., 2023).

3.3 S3 and R&I funding (ERDF dataset)

S3 priorities form the guiding foundation for the allocation of the European Regional Development Fund (ERDF). Bachtrögler et al. (2021) created a structured and comprehensive dataset compiling most of the projects co-funded by the ERDF (95% of the official commitments reported on the EC's Open Data Platform). The dataset contains projects from the 27 EU Member States and the United Kingdom over the programming period of 2014-2020 that were reported in lists of operations by a certain cut-off date (mostly end of 2020 or beginning of 2021).

For each project in this dataset, the following pieces of information are relevant: the NUTS2 region in which the project was carried out, the starting date of the project, a free text description of the project, a dummy variable indicating whether the project is R&I-related[4], and the actual ERDF co-funding amount. By text mining the free text descriptions[5], 42.4% of

[4] Since this study focusses on innovation, only R&I-related projects are included in the analysis (24.5% of the total ERDF co-funding amount).

[5] See *Appendix 7.2* for a detailed description of the text mining procedure.

the R&I projects could be matched to one or more technological domains[6]. On average, each of these projects were associated with 3.4 domains.

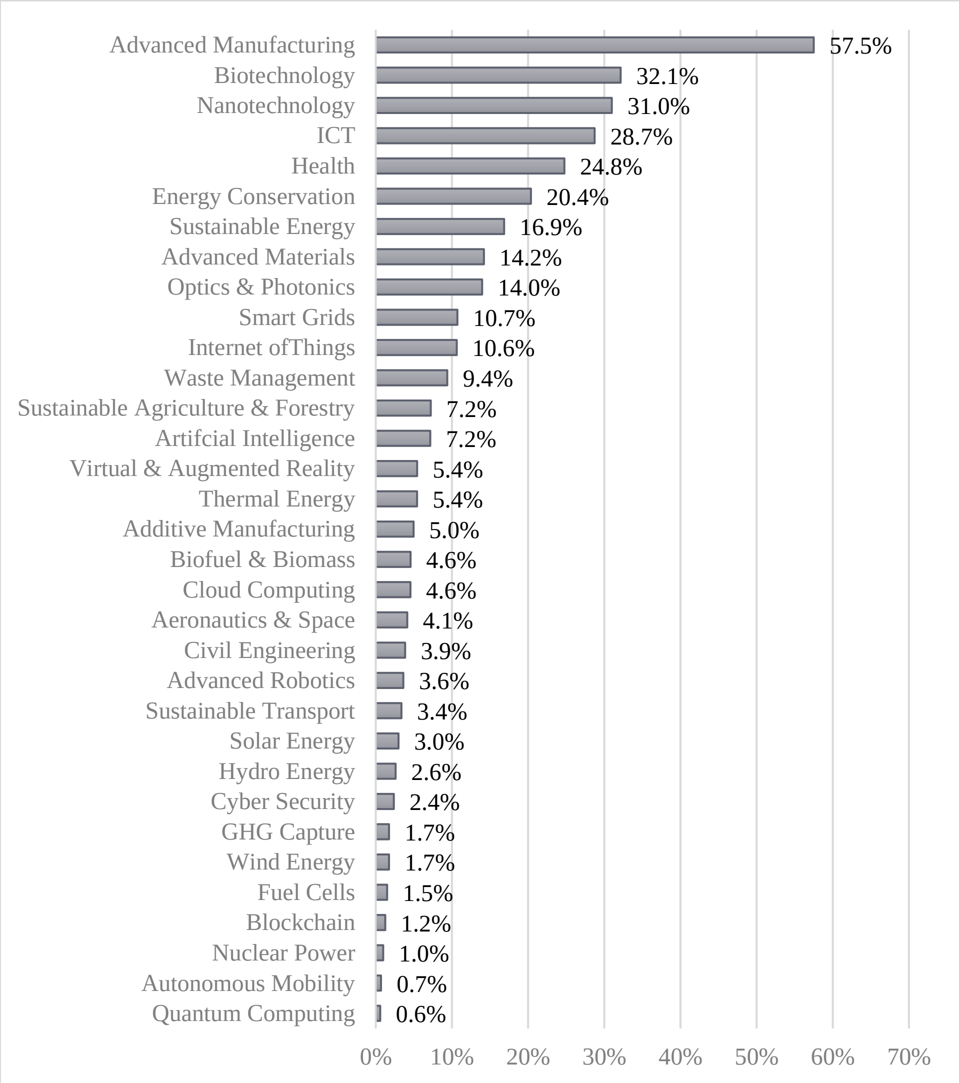

Figure 3. ERDF R&I projects per technological domain. Note: projects can be matched to more than one domain, so percentages do not add up to 100%.

Figure 3 shows the distribution of ERDF funded projects among the technological domains. The technological domains with the highest shares a generally also the domains that have a broad definition, therefore relatively many projects are related to those domains. In general, the domains with the lowest shares are either relatively new (e.g., blockchain) or more complex

[6] The remaining projects were excluded from the analysis.

(e.g., quantum computing) technologies. This is probably because these technologies are harder to imitate, therefore only a few firms work on R&I projects related to these technologies.

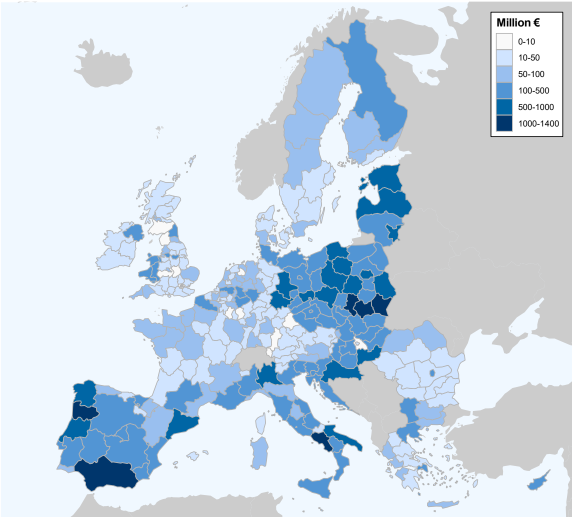

Figure 4a. Total ERDF funding for R&I projects per NUTS2 between 2014-2020.

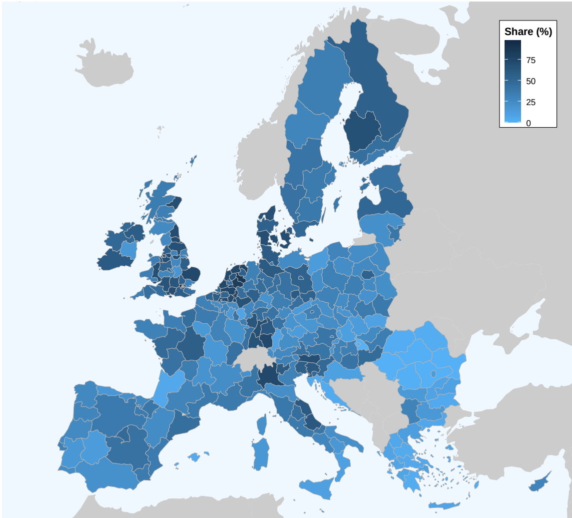

Figure 4b. Share of Total ERDF Expenditure dedicated to R&I projects between 2014-2020

As depicted in *Figure 4a*, the total amount of ERDF funding between 2014-2020 varies significantly among regions. Especially Southern and Eastern European regions (with the exception of Greece, Romania and Bulgaria) have a relatively high amount of R&I funding. *Figure 4b* puts this amount in the context of the total ERDF budget per NUTS2 region. Strikingly, the regions with high R&I expenditure in *Figure 4a*, are generally not the regions with a high share of their ERDF budget dedicated to R&I projects in *Figure 4b*. Most Dutch and British regions have a relatively high share of their total ERDF budget dedicated to R&I, but simultaneously have a low absolute amount of R&I funding. A possible explanation could be that there are more R&I-related activities going on in more developed regions, and so it is easier to find R&I-related projects to fund.

3.4 Regional technological capabilities (REGPAT dataset)

The OECD's REGPAT database is used to measure each region's technological capabilities. This dataset contains all patent applications to the European Patent Office between 1977 and 2020. The patent applications are regionalized at the NUTS2 level using the address of their inventor(s) and categorized according to this study's technological taxonomy (*Figure 2*) using their Cooperative Patent Classification (CPC) codes.

A region's technological capabilities are measured by the computation of a region's Relative Technological Advantage (RTA) in each technological domain and the degree of relatedness of each domain to the rest of the region's technological profile, also known as the relatedness density (Boschma et al., 2015). Based on Hidalgo et al. (2007), region *r* has an RTA in technology *i* at time *t* if the share of patents in technology *i* in region *r* is greater than the share of patents in technology *i* in the entire sample. More formally, RTA = 1 if

$$\frac{patents_{i,r,t} / \sum_i patents_{i,r,t}}{\sum_r patents_{i,r,t} / \sum_r \sum_i patents_{i,r,t}} > 1$$

Then, following Hidalgo et al. (2007) and Boschma et al. (2015), the density around technology *i* in region *r* at time *t* can be computed using the relatedness of technology *i* to the technologies in which region *r* has an RTA at time *t*, divided by the sum of technological relatedness of technology *i* to all the other technologies in the entire sample:

$$Relatedness\ Density_{i,r,t} = \frac{\sum_{\epsilon r, j \neq i} \varphi_{ij}}{\sum_{j \neq i} \varphi_{ij}}$$

RTA as well as relatedness density were calculated using the *EconGeo* R package by Balland (2017).

4. RESULTS

4.1 S3 PRIORITIES AND TECHNOLOGICAL CAPABILITIES

This section aims to analyse how well S3 priorities reflect a region's technological profile. As discussed in the previous section, each region prioritizes 15.6 technological domains on average. As *Figure 5a* shows, this amount varies considerably among regions. These regional differences can partially be explained by the relation between the number of prioritized technological domains and the number of domains in which a region is specialized (i.e., having an RTA above 1), shown in *Figure 6*.

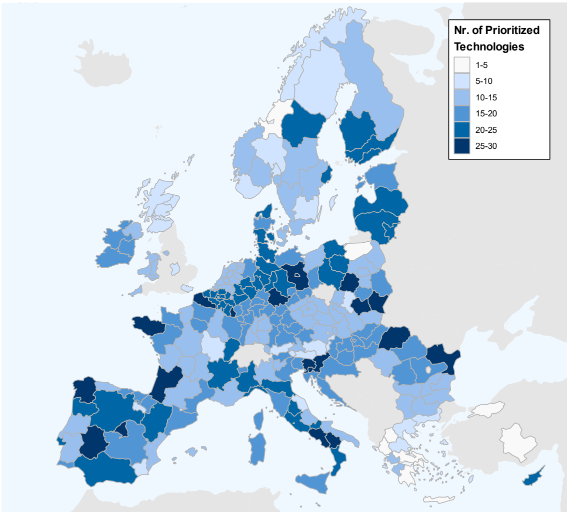

Figure 5a. Total number of prioritized technologies.

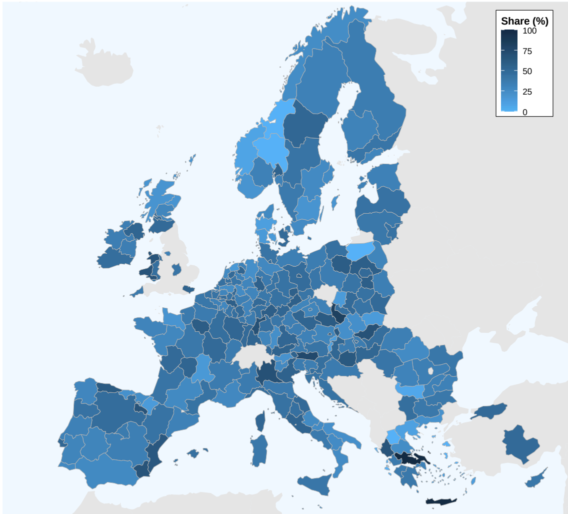

Figure 5b. Share of prioritized technologies in which a region has an RTA.

Strikingly, the relation shown in *Figure 6* seems to be stronger than shown in a similar graph for economic domains in a paper by Deegan et al. (2021, p. 509). Although this does not say anything decisive, it seems to be that regional prioritization choices for technological domains are generally more grounded than for economic domains. This view is supported by Kramer et al. (2021), who conclude that, in general, S3 priorities match better to a region's technological profile than to their economic profile. Moreover, *Figure 6* also shows that GDP does not seem to be related to the number of prioritized domains or the number of specializations. However, as expected, it appears that the more specializations a region has, the higher its share of prioritized domains in which it is specialized (see also *Figure 5b*).

Figure 5b maps each region's share of prioritized domains in which it has an RTA. This share averages 39.6% of prioritized domains, which is about the same as Marrocu et al. (2023) found for economic domains (43%). Kramer et al. (2021) also examined the correspondence between S3 priorities and technological domains and conclude that they match "relatively well" (p. 129). As they do not mention the average share, it is difficult to compare these findings one-to-one, but one could argue that a 39.6% match is not a "relatively well" match. This difference could be caused by their use of a less fine-grained taxonomy.

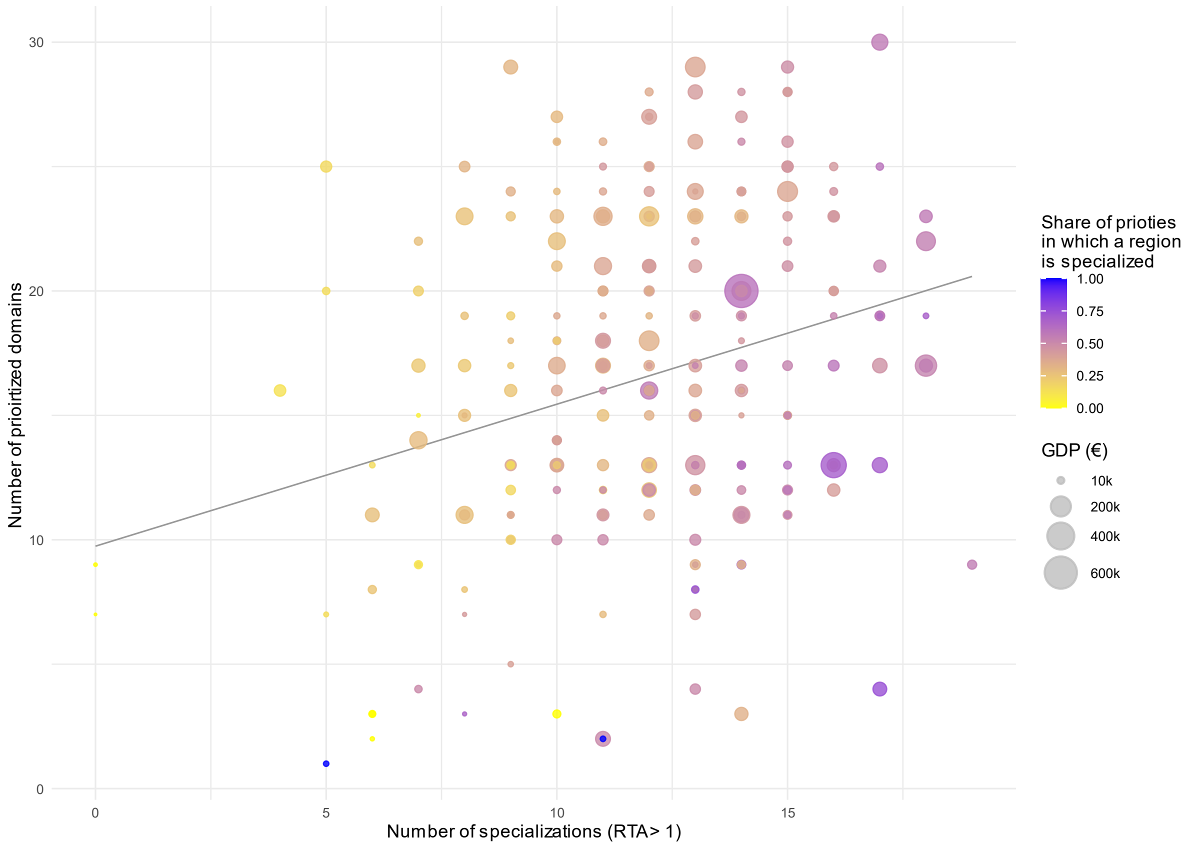

Figure 6. Relation between a region's total number of RTAs and their total number of prioritized technological domains.

Tables 3a-c reveal some more insights on the prioritization of technological domains. First of all, as presented in *Table 3a*, 16.44% of RTAs across EU regions are not prioritized. It could be that these were overlooked in the prioritization process, but it could also be that these are deliberately bypassed to avoid regional lock-ins. Second, the bottom row of each table shows the difference between prioritized and not prioritized technological domains in which regions are not specialized. The reason for prioritizing such an unacquired technology could be that the region aims to diversify into such a technological domain. As discussed before, a region is most likely to diversify into a new domain if it already possesses capabilities in related domains. Therefore, one would expect a large difference between the prioritized unacquired domains and unprioritized unacquired domains in *Table 3a* and *3b*. Although this is somewhat true, the difference seems to be marginal.

Table 3a. Distribution of technological domains.[7]

	Priority = 0	Priority = 1
RTA = 1	16.44%	19.51%
RTA = 0	34.23%	29.82%

Table 3b. Average relatedness density.

	Priority = 0	Priority = 1
RTA = 1	44.34	45.88
RTA = 0	31.80	33.37

Table 3c. Average RTA.

	Priority = 0	Priority = 1
RTA > 1	2.30	1.90
RTA < 1	0.46	0.52

Hypothesis 1 can be tested using a logit model in which the dependent variable represents the likelihood that region *r* prioritizes technological domain *i* explained by having an RTA in technology *i* (dummy) and the degree of relatedness of technology *i* to the region's technological portfolio. Hence, the basic equation to be estimated is written as follows:

$$Priority_{i,r,t} = \beta_0 + \beta_1 RTA_{i,r,t} + \beta_2 Relatedness\ Density_{i,r,t} + \varepsilon_{i,r,t}$$

where $Priority_{i,r,t} = 1$ if technology *i* is prioritized by region *r* in year *t*. *Table 4* shows that both variables stay significant, even when adding control variables and fixed effects for regions and years. The results support *Hypothesis 1*. Although regions tend to prioritize large sets of technological domains, they seem to consider their own technological profile when designing S3 priorities.

[7] There are 33 technologies for 256 NUTS2 regions, so 100% is equal to 8448 region-technology combinations.

Table 4. Correspondence between a region's S3 priorities and technological profile.

Dependent variable: Priority (dummy)	**Model 1 RTA & Rel. Density**	**Model 2 Covariates**	**Model 3 Fixed Effects**
Constant	-0.013164	0.010565	0.192831
	(0.020172)	(0.021330)	(0.181486)
RTA (dummy)	0.066404**	0.068480**	0.078054**
	(0.023442)	(0.024887)	(0.029880)
Relatedness density	0.162268***	0.142144***	0.140523**
	(0.023624)	(0.026549)	(0.051783)
GDP		-0.010628	0.006776
		(0.028277)	(0.075874)
Population density		-0.081329***	-0.056569
		(0.022344)	(0.053871)
Population size (log)		0.250414***	0.258608***
		(0.027683)	(0.069210)
Pseudo R^2 (McFadden)	0.007441	0.019309	
Conditional R^2			0.202
Region F.E.	No	No	Yes
Year F.E.	No	No	Yes
Num. Observations	9933	9042	9042
Num. NUTS2 regions	255	230	230
Num. Years	7	7	7

****P < 0.001; **P < 0.01; *P < 0.05*

4.2 S3 PRIORITIES AND R&I FUNDING

Next, I examine to what extent regions spend their R&I funding in line with their S3 priorities. Overall, 76% of all ERDF funding for R&I is spent on projects that are associated with a region's prioritized technological domains (see *Table 5*). It is worth noting that this percentage only considers the projects that could be matched to a technological domain via the text mining procedure (45.6% of all projects). It is hard to say anything decisive about the unmatched ERDF projects, as the cause could be either the inadequacy of the text mining process or the inadequacy of the text descriptions of these projects. The latter could imply a poor alignment with S3 priorities.

Table 5. Distribution of ERDF R&I funding.

	Priority = 0	**Priority = 1**
RTA = 1	10%	36%
RTA = 0	14%	40%

Figure 7 displays the spatial distribution of the share of ERDF expenditure on R&I projects that was in line with a region's prioritized technological domains. Most regions spend the overall share of their ERDF R&I funding in line with S3 priorities, however, there are some exceptions with a relative low share (such as Helsinki, Stockholm, Northern Denmark, and Northern Romania). This is mainly because these regions developed their S3 priorities nearing

the end of the programming period, and so, already spent a great deal of their funding before defining priorities.

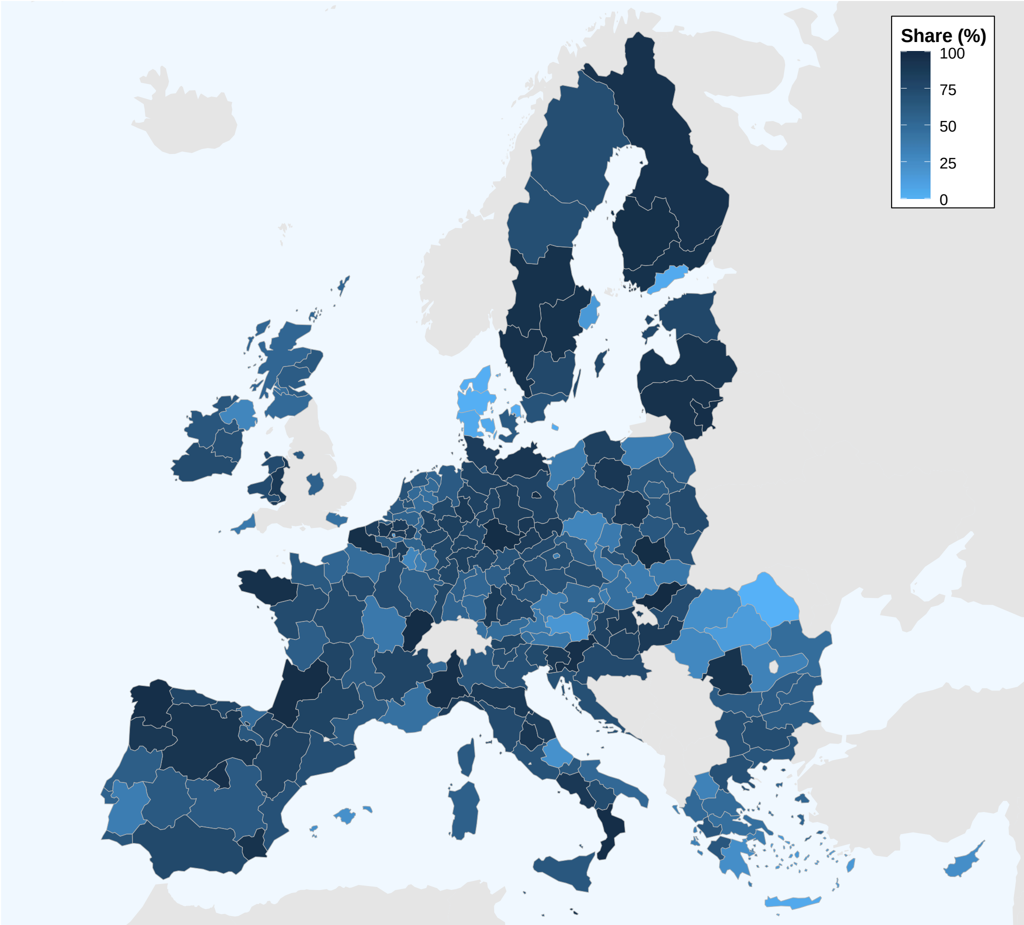

Figure 7. Share of ERDF expenditure on R&I in line with S3 priorities per NUTS2 region between 2014-2020.

To examine this correspondence in more detail, a logit model was created regressing whether technology i was funded by region r in year t over whether this technology was prioritized that year. The equation to be estimated can be written as follows:

$$Funded_{i,r,t} = \beta_0 + \beta_1 Priority_{i,r,t} + \varepsilon_{i,r,t}$$

where $Funded_{i,r,t} = 1$ if region r funded one or more projects associated with technology i in year t. *Table 6* shows there is a strong positive relation between prioritizing a technological domain and funding it. The variable priority maintains significant when adding fixed effects for countries, technologies and years.

Table 6. Correspondence between a region's ERDF spending and S3 priorities.

Dependent variable: ERDF funding (dummy)	**Model 1 Priority**	**Model 2 Covariates**	**Model 3 Fixed Effects**
Intercept	-0.663244***	-0.606078***	-1.961176
	(0.007987)	(0.009590)	(1.222037)
Priority	0.569061***	0.648513***	0.152211***
	(0.007750)	(0.009382)	(0.016811)
GDP		-0.053835***	-0.500360**
		(0.012599)	(0.154176)
Population density		0.007730	-0.151395
		(0.009675)	(0.138730)
Population size (log)		0.370779***	1.012923***
		(0.012924)	(0.171620)
Pseudo R^2 (McFadden)	0.056449	0.092238	
Conditional R^2			0.842
Region F.E.	No	No	Yes
Technology F.E.	No	No	Yes
Year F.E.	No	No	Yes
Num. Obs.	75999	54714	54714
Num. NUTS2 regions	256	238	238
Num. Technologies	33	33	33
Num. Years	9	7	7

$^{***}P < 0.001$; $^{**}P < 0.01$; $^{*}P < 0.05$

Although the results represented above suggest a confirmation of *Hypothesis 2*, a critical nuance should be made. As *Figure 8* evidently shows, there is a strong positive relation between each region's total number of prioritized technological domains and their share of R&I funding spent in line with those prioritized domains. Therefore, the high alignment between S3 priorities and R&I expenditure is mainly caused by the lack of selectivity in regional S3 strategies. For instance, the region of Brittany in the northeast corner of France, spent 95% of their funding on R&I projects that are related to their prioritized technological domains, which seems to support *Hypothesis 2*. However, of the 33 technological domains that are considered in this paper, they prioritized 30. Therefore, funding projects that are in line with their S3 strategy is relatively easy. While this may be self-evident, it is important to keep this in mind when interpreting these results. Especially when considering that most regions are not very selective in prioritizing technological domains (see *Figure 5a*).

So, although the results indicate a confirmation of *Hypothesis 2*, the high correspondence between S3 priorities and ERDF R&I funding is mainly a result of the overall large number of prioritized domains.

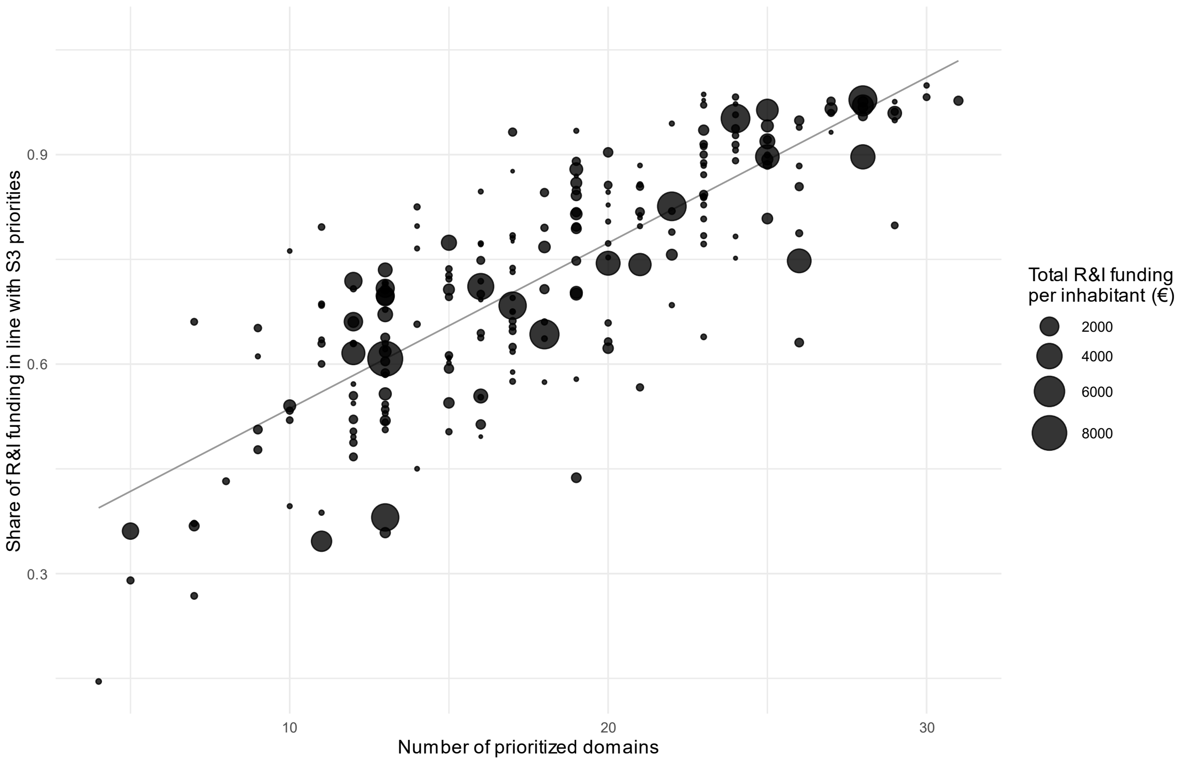

Figure 8. Relation between the number of prioritized technological domains and the share of R&I funding in line with those priorities per NUTS2 region.

4.3 R&I FUNDING AND KNOWLEDGE PRODUCTION

A knowledge production model was created to test the last hypothesis. The econometric model estimates the relation between a region's R&I expenditure in a technological domain and new patent applications related to that domain filed afterwards. To incorporate a sufficient time lag between R&I expenditure and patent applications, I use a 2-year lag between the starting year of an R&I project and the year a patent application is first filed to the patent office.

Because I am also interested in how the lack of selectivity in S3 strategies affects the impact of R&I funding on knowledge production, I include an interaction term. Rather than using a region's total number of prioritized domains, I use a more direct measure of a region's selectivity, namely the share of a region's total funding in year t that is dedicated to technological domain i. The underlying logic is that the more a region concentrates its resources, the more effective these are. Additionally, whether region r has an RTA in technology i and the relatedness of that technology to the regional technological profile are also considered.

As common in research using patent data, the current panel dataset has a large number of observations with zero patents. To deal with this issue, I follow the method proposed by Burger et al. (2009) and create a zero-inflated negative binomial model. This overcomes the biases usually created by logarithmic transformation and deals with overdispersion and excess zeros. Both the negative binomial and the zero-inflated part of the model are estimated using the following econometric equation:

$$Patents_{i,r,t} = \beta_0 + \beta_1 Expenditure_{i,r,t-2} + \beta_2 Expenditure_{i,r,t-2} * Share\ Total\ Exp._{i,r,t-2} + \beta_3 Relatedness\ \ Density_{i,r,t-2} + \beta_4 RTA_{i,r,t-2} + \varepsilon_{i,r,t-2}$$

Table 7 shows the results for the different econometric models. *Expenditure* is significant in all models, but, with the addition of covariates in *Model 3*, its sign reverses in both parts of the model. This is caused by the relation between *GDP* and *R&I Expenditure* (see *Figure 9a-b* in the *Appendix*). In general, regions with a higher GDP have less ERDF expenditure but generate more patents. The regional fixed effects in *Model 4* control for this.

Table 7. Knowledge production model.

Dependent variable: Patents (2-year lag)	**M1: Expenditure & Share of Tot. Exp.**		**M2: Relatedness density & RTA**		**M3: Covariates**		**M4: Fixed Effects**	
	NB	**Logit**	**NB**	**Logit**	**NB**	**Logit**	**NB**	**Logit**
Constant	1.929^{***}	-7.743^{***}	1.772^{***}	-7.606^{***}	1.608^{***}	-5.486^{***}	0.399	-3.816^{***}
	(0.013)	(0.899)	(0.012)	(0.845)	(0.013)	(0.185)	(0.420)	(0.294)
Log(Expenditure + 1)	0.040^{**}	-5.877^{*}	0.046^{***}	-4.709^{*}	-0.166^{***}	0.066^{*}	0.324^{***}	-0.812^{**}
	(0.012)	(2.652)	(0.012)	(2.285)	(0.011)	(0.034)	(0.012)	(0.257)
Interaction term: Exp. (log) * Share of Total Expenditure	0.201^{***}	-30.613^{***}	0.196^{***}	-27.339^{***}	0.235^{***}	-0.154^{***}	0.136^{***}	-4.189^{***}
	(0.011)	(7.270)	(0.010)	(6.252)	(0.009)	(0.026)	(0.006)	(0.782)
Relatedness density			0.223^{***}	-0.087	-0.112^{***}	-0.488^{***}	0.144^{***}	-0.544^{***}
			(0.012)	(0.189)	(0.013)	(0.036)	(0.023)	(0.129)
RTA (dummy)			0.402^{***}	-0.060	0.491^{***}	-0.019	0.352^{***}	0.073
			(0.012)	(0.162)	(0.012)	(0.032)	(0.013)	(0.105)
GDP					1.045^{***}	-10.324^{***}	0.535^{***}	-5.531^{***}
					(0.026)	(0.319)	(0.088)	(0.602)
Population size (log)					-0.110^{***}	0.269^{***}	0.621^{***}	0.582^{***}
					(0.024)	(0.035)	(0.101)	(0.163)
Population density					0.031	0.085	0.215	0.207
					(0.016)	(0.091)	(0.138)	(0.327)
Log(Theta)	-1.688^{***}		-1.604^{***}		-0.992^{***}			
	(0.008)		(0.008)		(0.011)			
AIC	230351.821		227479.890		177172.448		160134.896	
Log likelihood	-115168.911		-113728.945		-88569.224		-80046.448	
Num. Obs.	48180		48180		40425		40425	
Region F.E.	No		No		No		Yes	
Year F.E.	No		No		No		Yes	
Num. NUTS2 regions	289		289		242		242	
Num. Years	5		5		5		5	

$^{***}P < 0.001$; $^{**}P < 0.01$; $^{*}P < 0.05$

The interaction term is stable in all models and has a relatively large estimate (all variables are mean-centred), especially in the zero-inflated models. This means that the more expenditure is concentrated on a certain technological domain, the more effective it is in generating new knowledge. Moreover, *RTA* and *Relatedness density* are both significant (*RTA* only in the negative binomial model), indicating the importance of the consideration of both aspects in defining S3 priorities.

Overall, the model shows there is a relation between regional R&I funding of a certain technology and a region's knowledge production in that same technology 2 years later. This relation seems to become stronger when funding is concentrated on a few domains. Hence, these findings suggest a confirmation of *Hypothesis 3*.

5. Discussion and conclusions

This paper analysed three aspects of S3's first programming period: the correspondence between S3 priorities and regional capabilities, the alignment between S3 priorities and R&I funding, and the relation between R&I funding and knowledge production. A recurring issue throughout the analysis is the lack of selectivity in both S3 priorities and regional funding decisions. Although prioritized technological domains reflect to some extent regional technological profiles, there is still much room for improvement. The broadly defined priority areas target numerous technological domains often unrelated to a region's technological capabilities. Moreover, R&I funding seems to be well aligned with prioritized domains, but this is mainly caused by the large number of prioritized domains. Nevertheless, there is a significant and positive relation between R&I funding and knowledge production. However, this relation becomes stronger when funding is spent more concentrated as opposed to spread out over many domains. Therewith, unveiling the importance of selectivity for the effectiveness of R&I funding.

Still, being selective is not enough. As the analysis shows, a region's critical mass in a certain domain and its relatedness to a region's profile are also determinants of a region's knowledge production. Hence, regions should not only be selective, but also select those domains with the greatest growth potential. Therefore, selecting and defining priority areas should be substantiated by a sound analytical base. The extensive smart specialization literature has proposed several analytical tools and quantifiable concepts (like relatedness density and complexity) that could guide regions in this prioritization process. This paper and several others have shown that these concepts are only partially used by regional authorities. Providing regional policymakers with these analytical tools and the necessary data could help them to make more informed decisions.

Naturally, this study has some limitations. Although S3 tries to stimulate general regional development, this paper focusses solely on the technological dimension, more specifically technological capabilities captured in patent data. It is important to stress that patents do not reflect all technological capabilities and capture only a specific type of regional development. Therefore, this study should be complemented with research investigating other dimensions, as extensively discussed in the *Theoretical Background*.

Another, more methodological, limitation originates from the text mining procedure. Only 42.4% of the ERDF projects could be matched to one or more technological domains, which challenges the accuracy of the analysis of *Hypotheses 2* and *3*. This insufficiency could be caused by an inadequacy of the project descriptions or an inadequacy of the text mining process (see *Appendix 7.2* for more details). The former is difficult to solve, but there are several things future research can do to overcome the latter. For instance, by generating more extensive keyword lists for each technology, though this entails the risk of including keywords that are either too general or relate to several technological domains at once. Therefore, the addition of new keywords should always be complemented with manually scrutinizing the text mining results. Another option would be to use different, more advanced text mining techniques (for an overview see Talib et al. (2016) and Tandel et al. (2019)).

Although this study is the first to shed light on S3's short-term impact on regional knowledge production, its structural impact is yet to be revealed. In a similar vein as the current work, future research could focus on S3's long-term impact on knowledge production, but its impact on other dynamics, such as technological diversification, complexity, economic growth, or inter-regional collaborations, is important too.

On top of that, the current paper presents new evidence on the importance of selectivity in S3 strategies and the concentration of public resources on a few intervention areas. However, more analysis is still needed to explore in more detail what the most ideal composition of intervention areas would be. This probably depends on regional characteristics, but also on certain features of the targeted technological domains, such as their complexity or maturity.

Although this paper has shown the need for more selective S3 strategies, it is still unclear why regions tend to define large sets of broadly defined priority areas. While a few scholars have made some speculations about the underlying rationale (Iacobucci & Guzzini, 2016; Di Cataldo et al., 2021), there is still no substantiated understanding of why certain S3 principles are circumvented in the policy's implementation. Therefore, a natural progression of this work would be to investigate the regional implementation of S3 more closely to understand why regions tend to be unselective.

6. References

Asheim, B., Grillitsch, M., & Trippl, M. (2017). Smart Specialization as an Innovation-Driven Strategy for Economic Diversification: Examples From Scandinavian Regions. *Elsevier eBooks*, 73–97. https://doi.org/10.1016/b978-0-12-804137-6.00004-8

Antonelli, C., Crespi, F., & Quatraro, F. (2022). Knowledge complexity and the mechanisms of knowledge generation and exploitation: *The European evidence. Research Policy*, *51*(8), 104081. https://doi.org/10.1016/j.respol.2020.104081

Bachtrögler, J., Arnold, E., Doussineau, M., and Reschenhofer, P., (2021). UPDATE: Dataset of projects co-funded by the ERDF during the multi-annual financial framework 2014-2020. Publications Office of the European Union, Luxembourg, 2021, ISBN 978-92-76-43735-2, https://doi.org/10.2760/440159, JRC125008.

Balland, P.A. (2017) Economic Geography in R: Introduction to the EconGeo Package, *Papers in Evolutionary Economic Geography*, *17* (09), 1-75. Retrieved from http://econ.geo.uu.nl/peeg/peeg1709.pdf

Balland, P., & Rigby, D. L. (2017). The Geography of Complex Knowledge. *Economic Geography*, *93*(1), 1–23. https://doi.org/10.1080/00130095.2016.1205947

Balland, P., Boschma, R., Crespo, J., & Rigby, D. L. (2019). Smart specialization policy in the European Union: relatedness, knowledge complexity and regional diversification. *Regional Studies*, *53*(9), 1252–1268 https://doi.org/10.1080/00343404.2018.1437900

Balland, P., Jara-Figueroa, C., Petralia, S., Steijn, M., Rigby, D. L., & Hidalgo, C. A. (2020). Complex economic activities concentrate in large cities. *Nature Human Behaviour*, *4*(3), 248–254. https://doi.org/10.1038/s41562-019-0803-3

Boschma, R., & Wenting, R. (2007). The spatial evolution of the British automobile industry: Does location matter? *Industrial and Corporate Change*, *16*(2), 213–238. https://doi.org/10.1093/icc/dtm004

Boschma, R. (2014). Constructing Regional Advantage and Smart Specialisation: Comparison of Two European Policy Concepts. *SR Scienze regionali*, *1*, 51–68. https://doi.org/10.3280/scre2014-001004

Boschma, R., Balland, P., & 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. https://doi.org/10.1093/icc/dtu012

Burger, M., Van Oort, F., & Linders, G. (2009). On the Specification of the Gravity Model of Trade: Zeros, Excess Zeros and Zero-inflated Estimation. *Spatial Economic Analysis*, *4*(2), 167–190. https://doi.org/10.1080/17421770902834327

Buyukyazici, D. (2022). Skills for Smart Specialization: Relatedness, Complexity and Evaluation of Priorities. *Papers in Evolutionary Economic Geography*, *No. 22.18*. Utrecht: Department of Economic Geography, University of Utrecht. http://econ.geo.uu.nl/peeg/peeg2218.pdf

Camagni, R., & Capello, R. (2013). Regional Innovation Patterns and the EU Regional Policy Reform: Toward Smart Innovation Policies. *Growth and Change*, *44*(2), 355–389. https://doi.org/10.1111/grow.12012

Capello, R., & Kroll, H. (2016). From theory to practice in smart specialization strategy: emerging limits and possible future trajectories. *European Planning Studies*, *24*(8), 1393–1406. https://doi.org/10.1080/09654313.2016.1156058

Crescenzi, R., de Blasio, G., & Giua, M. (2020). Cohesion Policy incentives for collaborative industrial research: evaluation of a Smart Specialisation forerunner programme. *Regional Studies*, *54*(10), 1341–1353. https://doi.org/10.1080/00343404.2018.1502422

D'Adda, D., Iacobucci, D., & Perugini, F. (2021). Smart Specialisation Strategy in practice: have regions changed the allocation of Structural Funds? *Regional Studies*, *56*(1), 155–170. https://doi.org/10.1080/00343404.2021.1890326

Deegan, J., Broekel, T., & Fitjar, R. D. (2021). Searching through the Haystack: The Relatedness and Complexity of Priorities in Smart Specialization Strategies. *Economic Geography*, *97*(5), 497–520. https://doi.org/10.1080/00130095.2021.1967739

Essletzbichler, J. (2013). Relatedness, Industrial Branching and Technological Cohesion in US Metropolitan Areas. *Regional Studies, 49*(5), 752–766. https://doi.org/10.1080/00343404.2013.806793

European Commission (2018). *Eye@RIS3 - Smart Specialisation Platform*. Smart Specialisation Platform. Retrieved October 13, 2022, from https://s3platform.jrc.ec.europa.eu/map/-/eye3/y/2014-2020

European Commission (2022, December 31). *European Regional Development Fund (ERDF)*. Cohesion Open Data Platform. Retrieved March 29, 2023, from https://cohesiondata.ec.europa.eu/funds/erdf/14-20

Farinha, L., Lopes, J. T., Sebastião, J., Ferreira, J. J., Oliveira, J. L., & Silveira, P. C. L. (2021). How do stakeholders evaluate smart specialization policies defined for their regions? *Competitiveness review*. https://doi.org/10.1108/cr-12-2019-0149

Fratesi U., Gianelle, C. & Guzzo, F. (2021). Assessing Smart Specialisation: Policy Implementation Measures, EUR 30758 EN, Publications Office of the European Union, Luxembourg, 2021, ISBN 978-92-76-40034-9, https://doi.org/10.2760/71769, JRC123821

Foray, D., David, P. A., & Hall, B. (2009). *Smart specialization – The Concept.* (Knowledge Economists Policy Brief No. 9, June). Brussels: European Commission. Retrieved from https://ec.europa.eu/invest-in-research/pdf/download_en/kfg_policy_brief_no9.pdf?11111

Foray, D., David, P. A., & Hall, B. H. (2011). Smart specialization. From academic idea to political instrument, the surprising career of a concept and the difficulties involved in its implementation (MTEI Working Paper, November). Lausanne: MTEI.

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

Specialisations (RIS 3), Brussels. Retrieved from https://s3platform.jrc.ec.europa.eu/en/w/guide-on-research-and-innovation-strategies-for-smart-specialisation-ris3-guide-

Foray, D. (2014). *Smart Specialisation: Opportunities and Challenges for Regional Innovation Policy*. Routledge. https://doi.org/10.4324/9781315773063

Grillitsch, M. (2016). Institutions, smart specialisation dynamics and policy. *Environment and Planning C-government and Policy*, *34*(1), 22–37. https://doi.org/10.1177/0263774x15614694

Gianelle, C., Guzzo, F., & Mieszkowski, K. (2020a). Smart Specialisation: what gets lost in translation from concept to practice? *Regional Studies*, *54*(10), 1377–1388. https://doi.org/10.1080/00343404.2019.1607970

Gianelle, C., Kyriakou, D., McCann, P., & Morgan, K. (2020b). Smart Specialisation on the move: reflections on six years of implementation and prospects for the future. *Regional Studies*, *54*(10), 1323–1327. https://doi.org/10.1080/00343404.2020.1817364

Hassink, R., & Gong, H. (2019). Six critical questions about smart specialization. *European Planning Studies*, *27*(10), 2049–2065. https://doi.org/10.1080/09654313.2019.1650898

Hidalgo, C., Klinger, B., Barabási, A., & Hausmann, R. (2007). The Product Space Conditions the Development of Nations. *Science*, *317*(5837), 482–487. https://doi.org/10.1126/science.1144581

Hidalgo, C. A., Balland, P., Boschma, R., Delgado, M., Feldman, M. P., Frenken, K., Glaeser, E. L., He, C., Kogler, D. F., Morrison, A., Neffke, F., Rigby, D. L., Stern, S., Zheng, S., & Zhu, S. (2018). The Principle of Relatedness. *Springer eBooks*, 451–457. https://doi.org/10.1007/978-3-319-96661-8_46

Kogler, D. F., Rigby, D. L., & Tucker, I. (2013). Mapping Knowledge Space and Technological Relatedness in US Cities. *European Planning Studies*, *21*(9), 1374–1391. https://doi.org/10.1080/09654313.2012.755832

Kroll, H. (2015). Efforts to Implement Smart Specialization in Practice—Leading Unlike Horses to the Water. *European Planning Studies*, *23*(10), 2079–2098. https://doi.org/10.1080/09654313.2014.1003036

Kroll, H. (2019). Eye to eye with the innovation paradox: why smart specialization is no simple solution to policy design. *European Planning Studies*, *27*(5), 932–951. https://doi.org/10.1080/09654313.2019.1577363

Landabaso, M. (2014). Guest editorial on research and innovation strategies for smart specialisation in Europe. *European Journal of Innovation Management*, *17*(4), 378–389. https://doi.org/10.1108/ejim-08-2014-0093

Lopes, J. T., Farinha, L., Ferreira, J. J., & Silveira, P. C. L. (2018). Does regional VRIO model help policy-makers to assess the resources of a region? A stakeholder perception approach. *Land Use Policy*, *79*, 659–670. https://doi.org/10.1016/j.landusepol.2018.07.040

Lopes, J. T., Ferreira, J. J., & Farinha, L. (2019). Innovation strategies for smart specialisation (RIS3): Past, present and future research. *Growth and Change*, *50*(1), 38–68. https://doi.org/10.1111/grow.12268

Marinelli, E., & Perianez-Forte, I. (2017). Smart Specialisation at work: The entrepreneurial discovery as a continuous processS3. *S3 Working Paper Series*, *12/2017*. https://doi.org/10.2760/514714

Marrocu, E., Paci, R., Rigby, D. L., & Usai, S. (2022). Evaluating the implementation of Smart Specialisation policy. *Regional Studies*, *57*(1), 112–128. https://doi.org/10.1080/00343404.2022.2047915

McCann, P., & Ortega-Argilés, R. (2015). Smart Specialization, Regional Growth and Applications to European Union Cohesion Policy. *Regional Studies*, *49*(8), 1291–1302. https://doi.org/10.1080/00343404.2013.799769

McCann, P., & Ortega-Argilés, R. (2016). The early experience of smart specialization implementation in EU cohesion policy. *European Planning Studies*, *24*(8), 1407–1427. https://doi.org/10.1080/09654313.2016.1166177

Mewes, L., & Broekel, T. (2022). Technological complexity and economic growth of regions. *Research Policy*, *51*(8), 104156. https://doi.org/10.1016/j.respol.2020.104156

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*(3), 237–265. https://doi.org/10.1111/j.1944-8287.2011.01121.x

OECD (August, 2022) REGPAT Database.

Petralia, S., Balland, P. A., & Morrison, A. (2017). Climbing the ladder of technological development. *Research Policy*, *46*(5), 956–969. https://doi.org/10.1016/j.respol.2017.03.012

Pugh, R. (2018). Questioning the implementation of smart specialisation: Regional innovation policy and semi-autonomous regions. *Environment and Planning C: Politics and Space*, *36*(3), 530–547. https://doi.org/10.1177/2399654417717069

Rigby, D. L. (2013). Technological Relatedness and Knowledge Space: Entry and Exit of US Cities from Patent Classes. *Regional Studies*, *49*(11), 1922–1937. https://doi.org/10.1080/00343404.2013.854878

Rigby, D. L., Roesler, C., Kogler, D., Boschma, R., & Balland, P. A. (2022). Do EU regions benefit from Smart Specialisation principles? *Regional Studies*, *56*(12), 2058–2073. https://doi.org/10.1080/00343404.2022.2032628

Rodríguez-Pose, A., Di Cataldo, M., & Rainoldi, A. (2014). The Role of Government Institutions for Smart Specialisation and Regional Development. *RePEc: Research Papers in Economics*. https://publications.jrc.ec.europa.eu/repository/bitstream/JRC88935/2014-jrc_tech_rep_cov_templ_online_s3p_rodpose%20final.pdf

Romão, J. (2020). Variety, smart specialization and tourism competitiveness. *Sustainability (Switzerland)*, *12*(14), 1–13. https://doi.org/10.3390/su12145765

Santos, A. C., Edwards, J. E., & Neto, P. A. M. (2022). Does Smart Specialisation improve innovation subsidy effect on regional productivity? The Portuguese case. *European Planning Studies*, 1–22. https://doi.org/10.1080/09654313.2022.2073787

Schmoch, U. (2008). Concept of a Technology Classification for Country Comparisons - Final Report. Fraunhofer Institute for Systems and Innovation Research, Karlsruhe, Germany. https://www.wipo.int/ipstats/en/statistics/patents/pdf/wipo_ipc_technology.pdf

Sörvik, J. & Kleibrink, A. (2015). ‘Mapping Innovation Priorities and Specialisation Patterns in Europe’, JRC Technical Reports JRC95227.

Storper, M. (1997). *The Regional World: Territorial Development in a Global Economy*. Guilford Publications.

Tandel, S. S., Jamadar, A., & Dudugu, S. (2019). A Survey on Text Mining Techniques. *International Conference on Advanced Computing*. https://doi.org/10.1109/icaccs.2019.8728547

Talib, R., Abbas, G., Ayesha, S., & Fatima, F. (2016). Text Mining: Techniques, Applications and Issues. *International Journal of Advanced Computer Science and Applications*, *7*(11). https://doi.org/10.14569/ijacsa.2016.071153

Trippl, M., Zukauskaite, E., & Healy, A. (2020). Shaping smart specialization: the role of place-specific factors in advanced, intermediate and less-developed European regions. *Regional Studies*, *54*(10), 1328–1340. https://doi.org/10.1080/00343404.2019.1582763

7. Appendix

7.1 Regional and temporal dimension of S3 priority data

Most S3 strategies were developed at the NUTS2 level, however, some also at the NUTS1, NUTS3 or national level[8]. In order to make comparable analysis possible, all strategies were converted to the NUTS2 level. This entailed duplicating NUTS1 region strategies to its subordinate NUTS2 regions, and aggregating NUTS3 level strategies to their overarching NUTS2 region. The S3 strategies at the national level were excluded from the analysis, except for smaller countries where the national level is the same as the NUTS2 level, such as Luxembourg and Latvia.

The Eye@RIS3 dataset also includes the date on which the strategy was submitted to the S3 platform. This date was incorporated in the econometric analysis for both *Hypothesis 1* and *2*. This means that regional capabilities were measured in the year that a strategy was designed, and that the correspondence between S3 priorities and R&I funding is measured from the year a strategy was designed to the end of the programming period.

7.2 Text mining process

For each technological domain, a comprehensive list of associated keywords was created. This was done by manually selecting several glossary/terminology websites for each

[8] For a more detailed description of the varying territorial levels of S3 strategies, see Di Cataldo et al. (2021).

technology[9]. These websites contain key terms associated with a certain technology which were automatically scraped and then compiled per technological domain. Keywords that were too general were deleted. For instance, the term ‘kilowatt per hour’, which can be associated with several technologies that focus on the generation of energy. The keyword lists that resulted from this were then used to automatically text mine the ERDF project descriptions. The results of this matching process were then manually scrutinized to fine-tune the keywords list, whereafter the automatic text mining exercise was repeated. These last steps were reiterated several times until the matching results seemed satisfactory.

7.3 The relation between GDP, R&I funding, and knowledge production

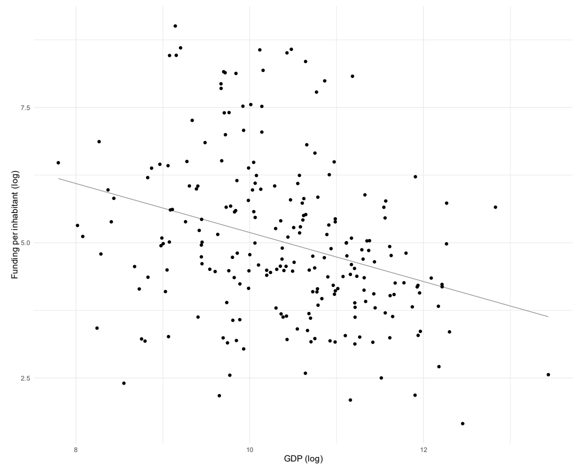

Figure 9a. The relation between GDP and ERDF R&I funding per capita. Note: both variables are log transformed.

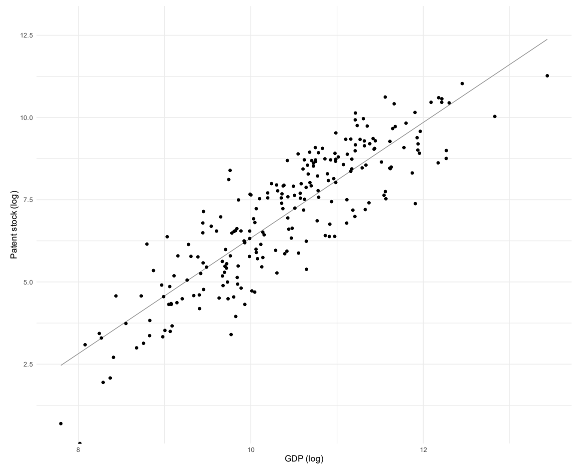

Figure 9b. The relation between GDP and regional patent stock. Note: both variables are log transformed.

[9] For example, https://en.wikipedia.org/wiki/Glossary_of_artificial_intelligence and https://www.expert.ai/glossary-of-ai-terms/ were, among others, used to gather keywords associated with Artificial Intelligence.