Papers in Evolutionary Economic Geography

20.49

Technological Complexity and Economic Growth of Regions

Lars Mewes & Tom Broekel

Utrecht University Human Geography and Planning

Technological Complexity and Economic Growth of Regions

Lars Mewes

Institute of Economic and Cultural Geography, Leibniz University Hannover, Germany
mewes@wigeo.uni-hannover.de

Tom Broekel

University of Stavanger Business School, Norway
Center for Regional and Innovation Economics, University of Bremen, Germany
tom.brokel@uis.no

Abstract

One the one hand, complex technologies offer substantial economic benefits, and on the other, they are difficult to invent and to imitate, and they refuse a fast dissemination. This two-sidedness motivates the idea that regions' competitive advantages and, in consequence, their economic growth, originate in their ability to produce and utilize complex technologies. However, the relationship between technological complexity and regional economic growth has rarely been empirically investigated. Here, we address this pressing issue by assessing the complexity of technological activities in 159 European NUTS 2 regions and relating it to their economic growth from 2000 to 2014. Our empirical results suggest that technological complexity is an important predictor of regional economic growth. A 10% increase in complexity is associated with a 0.45% GDP per capita growth. By showing that technological complexity is important for regional economic growth, our results fuel current policy debates about optimal regional policies such as the Smart Specialization strategy.

Keywords: *Knowledge Complexity, Technological Complexity, Regional Economic Growth, Patent Data*

JEL: *O10, O33, R11*

1 Introduction

"I remember thinking how comfortable it was, this division of labor which made it unnecessary for me to study fogs, winds, tides, and navigation, in order to visit my friend who lived across an arm of the sea. It was good that men should be specialists [...]. The peculiar knowledge of the pilot and captain sufficed for many thousands of people who knew no more of the sea and navigation than I knew" (London, 1904).

Over a hundred years ago, Humphrey van Weyden praised the benefits of specialization and division of labor aboard a small vessel in Jack London's famous novel, "The Sea Wolf." They allowed him to concentrate on those things that caught his interest and talents. Besides their effects on individual well-being, specialization and the division of labor increased productivity, generated economic surpluses and allowed for sustaining larger population sizes (Smith, 1776). The coordination and cooperation of specialists to utilize the large amounts of diverse knowledge is easier in larger and more densely populated areas (Becker and Murphy, 1992). In turn, such larger and more densely connected populations fuel further specialization and division of labor (Sveikauskas, 1975). This self-reinforcing process accelerated the richness and complexity of knowledge production over time (Kremer, 1993; Henrich, 2004). Hence, one implicit consequence of specialization and division of labor is the constantly increasing complexity of the world's knowledge (Aunger, 2010).

Knowledge, in general, represents a critical resource in today's knowledge economy (Lucas, 1988; Romer, 1990). However, not all knowledge is alike and equally valuable. More complex knowledge is argued to be a fundamental building block of competitive advantage and economic growth (Kogut and Zander, 1992; Hidalgo and Hausmann, 2009). Its economic relevance rests on the idea that complex knowledge is difficult to imitate, and that only few economic actors have the capabilities to produce it (Storper, 2010). Consequently, firms and economies with complex knowledge are likely to earn rents in the form of higher growth and wealth (Kogut and Zander, 1992; Teece et al., 1997; Hidalgo and Hausmann, 2009).

Until now, empirical evidence of the economic benefit of knowledge complexity is scarce and restricted to economic complexity as measured by the product portfolio of an economy (Hidalgo and Hausmann, 2009; Hausmann et al., 2013; Bahar et al., 2014). Production, however, is only one dimension of complexity, in which knowledge represents a critical resource for building competitive advantage. Technological know-how is complementary and similarly vital for economies' competitiveness and growth (Nelson and Winter, 1982; Lucas, 1988; Romer, 1990). Yet, the relation between technological complexity and economic growth is still unexplored.

In this article, we address this research gap by studying the relationship between technological complexity and economic growth at the level of 159 European *NUTS 2* regions between 2000 and 2014. We approximate technological activities by relying on patent documents (Fleming and Sorenson, 2001), and assess technological complexity with the recently developed measure of *Structural Diversity* (Broekel, 2019). The results of dynamic panel regressions confirm that technological complexity is a positive and robust predictor of economic growth in European regions.

Our study is structured as follows. Section 2 provides an overview of the theoretical and empirical literature on technological complexity. Section 3 presents the empirical data and our estimation approach. The empirical results are presented in Section 4. Based on our findings, Section 5 concludes that technological complexity can be considered an important dimension of regional knowledge production capable of informing current policy programs such as the Smart Specialization strategy of the European Union.

2 Theoretical background and literature overview

Knowledge production is a fundamental source of long-term economic growth (Kuznets, 1962; Nelson and Winter, 1982; Romer, 1990; Grossman and Helpman, 1991; Aghion and Howitt, 1998), which helps in understanding the uneven growth patterns of regions (Glaeser et al., 1992; Fagerberg et al., 1997; Henderson et al., 2001). Knowledge accumulates over time in - and adheres to - certain locations, leading to a strong spatial concentration of knowledge in regions (Feldman, 1994). One important reason for the spatial concentration of knowledge is the sensitivity of knowledge spillovers to geographic distance, as this limits the spatial diffusion of knowledge and contributes to its geographic concentration (Jaffe et al., 1993; Markusen, 1996). Crucially, the degree of spatial concentration varies significantly between knowledge domains (Breschi and Malerba, 1997). While researchers in the past have highlighted the role of tacit knowledge in this context (Lawson and Lorenz, 1999; Gertler, 2003), the complexity of knowledge has increasingly been focused on as one crucial dimension that explains the varying spatial concentration of knowledge domains (Hidalgo and Hausmann, 2009).

In contrast to the intensity of the discussion on the complexity of knowledge and its economic relevance in the literature, there is (still) no common definition of knowledge complexity. Yet, there seems to be a consensus on a number of its basic features. To Winter (1987, p. 177), the complexity of knowledge is *"the amount of information required to characterize the item of knowledge in question."* Zander and Kogut (1995) rely on a similar understanding, which focuses on the diversity of knowledge combination. Accordingly, knowledge *"is more complex when it draws upon distinct and multiple kinds of components"* (Zander and Kogut, 1995, p. 79). Kauffman (1993) defines complexity in a related manner, as the interaction between size and interdependence of components. This builds on Simon's (1962) description of complex systems. For him, complexity is *"made up of a large number of parts that interact in a nonsimple way"* (Simon, 1962, p. 468). Interestingly, there is a similarity to Polanyi's (1966) notion of tacitness, where the more information - e.g., a diverse range of combinations, interdependencies, and competences - a system entails, the more difficult becomes communication and codification.

While knowledge complexity generally includes all types of knowledge, its fundamental features also apply to technologies. Technologies are often described as compositions of multiple components that are combined to fulfill a specific purpose (Usher, 1954; Hargadon, 2003; Arthur, 2009). The number of components, their intensity of combination, and how they are combined are seen as primary determinants of their complexity (Fleming and Sorenson, 2001; Broekel, 2019). While these components include *knowledge bits* and *matter* (Fleming and Sorenson, 2004), we primarily approach them from a knowledge perspective. Hence, we will use both terms, knowledge complexity and technological complexity, interchangeably throughout the article[1].

Knowledge complexity has crucial effects on knowledge creation in an economy. Complexity is one important qualitative dimension of knowledge that determines the cost and time of knowledge imitation. Errors in imitation tend to become more frequent with growing complexity, suggesting that imitation is not a promising strategy in complex knowledge domains (Rivkin, 2000). Hence, complex knowledge is less likely to spillover to competitors. Yayavaram and Chen (2015) demonstrate that the acquisition of new and complex knowledge in innovation processes impedes learning and hurts innovation outcomes. Consequently, complex knowledge represents an entry barrier, as it is more difficult to learn and to copy.

The ability to learn and acquire complex knowledge is therefore argued to be more valuable and to translate into higher economic rents than knowledge that can be easily acquired (Winter,

[1]Note that economic complexity is a another term used in the literature (Hidalgo and Hausmann, 2009), and primarily refers to the empirical assessment of the complexity of economic activities using data on export products.

1987; Kogut and Zander, 1992; Zander and Kogut, 1995; Teece et al., 1997; Storper, 2010). As complex knowledge represents a critical resource, economic actors can build a competitive advantage based on complex knowledge, providing them with profound growth potentials and access to quasi-monopolistic rents (Teece, 1977; Kogut and Zander, 1992; Zander and Kogut, 1995; Teece et al., 1997; Rivkin, 2000; McEvily and Chakravarthy, 2002; Sorenson et al., 2006). Indeed, empirical insights back this argument. For instance, Fleming and Sorenson (2001) show that more complex inventions receive more citations, indicating that complex inventions are technologically more valuable than simpler inventions. Their benefits are also shown to be better appropriated by their inventors (Sorenson et al., 2006).

Crucially, geographic proximity plays an important role in the creation and diffusion of complex knowledge. It is widely accepted and empirically confirmed that geographic proximity facilitates interactions and engagement in networks (Becker et al., 1999; Boschma, 2005; Breschi and Lissoni, 2009). Thereby, geographic proximity stimulates the interactive learning required for the creation of complex knowledge. In addition, it eases its exchange by allowing for easier and quicker feedback, spontaneous interactions of heterogeneous actors, and more efficient communication (Malmberg and Power, 2005). Empirical confirmation for these arguments are provided by Balland and Rigby (2017). These authors find that complex technologies diffuse slower in space than simple ones. Broekel (2019) adds to this by notion, finding competences in complex technologies as being concentrated in space. Importantly, such concentrations are not randomly distributed in space, but rather seem to be increasingly be found in urban agglomerations (Balland et al., 2020).

The geographic concentration of complex knowledge suggests that not every region has the capabilities to produce it. This observation paired with the argument that highly complex knowledge is economically more valuable than less complex knowledge raises the crucial question: do regions benefit from activities in complex technologies? Asked differently, do complex technological activities facilitate the economic growth of regions?

Most existing studies have focused on economic complexity and the country level. In their seminal paper, Hidalgo and Hausmann (2009) introduce the *Economic Complexity Index (ECI)* to approximate the complexity of countries' economic activities based on their production capabilities. The ECI builds on the spatial distribution of export products across countries. In this framework, products (and the knowledge underlying their production) exported by few and most diversified economies are assumed to be more complex. On this basis, the authors show empirically that countries with greater economic complexity are characterized by higher levels of GDP per capita and experience higher short-term GDP growth. Subsequent studies have supported the findings of Hidalgo and Hausmann (2009) that economic complexity matters for countries' economic growth (Ferrarini and Scaramozzino, 2016; Stojkoski et al., 2016). The production of goods and services is only one dimension in which countries and regions compete; another important one is technological know-how. However, technological complexity is distinct from economic complexity, as the latter also considers (countries') capabilities in terms of institutions, infrastructure, and labor skills (Hidalgo and Hausmann, 2009). Accordingly, insights based on economic complexity cannot be applied directly to the context of technological capabilities. Moreover, existing empirical evidence focuses on the national level, ignoring the substantial variations of (technological) capabilities at the sub-national level (Balland and Rigby, 2017; Balland et al., 2020). A notable exception is the recent study by Antonelli et al. (2020), wherein the authors focus on the regional level and link (knowledge) complexity to productivity. However, their empirical results suggest that while complexity positively relates to the growth of technological knowledge in regions, it appears to have a negative effect on regional productivity growth. This somewhat contrasts with the theoretical expectations outlined above. Consequently, there is a need to shed further light on this issue. Such need motivates the present paper, which aims to empirically test the relationship between technological complexity and regional economic growth.

3 Materials and methods

Our unit of analysis are NUTS 2 regions in Europe, for which we collected a rich set of variables for all years between 2000 and 2014. We chose NUTS 2 regions primarily for reasons of data availability. Clearly, labor market regions would be more appropriate to capture the regional dimension of innovation processes. However, there is no common definition across Europe, and many empirical variables are not available at other levels. The final sample is composed of 166 unique regions observed over 15 years. In a common manner, we approximated economic growth by the annual change in GDP per capita, which we obtained from *Eurostat*. On this basis, we defined our dependent variable as GDP per capita growth (Power Purchasing Standards) in region r and year t.

3.1 Technological complexity

3.1.1 Measuring technological complexity

We used patent data of the *OECD REGPAT Database* (March 2018 version), which covers patent applications to the *European Patent Office (EPO)*, as an indicator of technologies. Although patents come with several disadvantages, they are nevertheless widely used in empirical research on technological knowledge production (Griliches, 1990). This is mainly because patents are the only large-scale data source providing such detailed information about technological knowledge.

Calculating technological complexity is not a straightforward task, as there is no established method thus far. The ECI by Hidalgo and Hausmann (2009) seems to be the most prominent approach in today's literature. However, it was developed to assess the economic complexity of countries based on their export portfolios. While the ECI has been used to approximate technological complexity using information on patent activities of countries and regions (Balland and Rigby, 2017; Petralia et al., 2017; Antonelli et al., 2020), such applications face a number of issues. For instance, the obtained knowledge complexity index (KCI) is based on the spatial distribution of technologies, which may create endogeneity issues in spatial research. The spatial distribution of technologies is shaped by many factors (e.g., natural conditions, institutions, infrastructure), of which complexity is but one. Most importantly, many of the KCI's empirical characteristics do not reflect what is generally expected of technological complexity (Tacchella et al., 2012; Broekel, 2019).

We therefore relied on the measure of *Structural Diversity* that was recently developed by Broekel (2019). Besides its empirical advantages, the measure also resembles more closely the theoretical foundations of technological complexity, as presented in the previous section[2]. Structural Diversity relies on information theory and assesses the diversity of how knowledge components of a technology are combined. It rests on the idea that new knowledge and technologies are developed through (re-)combinatorial processes (Hargadon, 2003; Arthur, 2009). Consequently, they can be described as so-called combinatorial networks. These networks consist of nodes, representing the different knowledge components, and links, indicating their combinations. For instance, a chair can be seen as a combination of four chair legs, one seat, and one backrest. The four legs are identical from a knowledge perspective and hence represent just one distinct knowledge component implying that the chair consists of three distinct (knowledge) components.

[2]We also used the KCI approach (Hidalgo and Hausmann, 2009; Balland and Rigby, 2017) to estimate technological complexity (see Section 4.2 for more details).

The idea of Structural Diversity is to measure the diversity of how these components are combined with each other. In case of the chair, all four legs are directly "combined" with the seat: they are not connected to each other nor to the backrest, and the backrest is also directly "combined" with the seat. Accordingly, the combinatorial network of the chair corresponds to a star-like network composed of one central component (the seat) and two peripheral components (legs and backrest). Crucially, this network has just one topology (a star), which implies that little information is required for its complete description, and it can thereby be regarded as a relatively simple (ordered) network.

As a contrasting example, consider the combinatorial network of a car. This network has more components and consequently consists of more nodes and links. However, this is not the primary reason that makes it more complex than the network of the chair. According to Broekel (2019), the network becomes more complex because it features a greater number of distinct topologies. Some of its components will be "combined" in a star-like manner (front, back, and side windows with the car body), while others may rather be connected in form of a "line": steering wheel to steering column to steering gear[3]. From an information theoretical perspective, a greater diversity of topologies (distinct combinatorial structures) implies that more information is required to describe the combinatorial network of a car than that of a chair. Due to its greater information content, the combinatorial network of the car qualifies as being more complex than that of the chair (Emmert-Streib and Dehmer, 2012). On this basis, Broekel (2019) argues that an index reflecting the diversity of distinct topologies in such networks is able to differentiate between simple and complex networks, or, as in the case of technologies' combinatorial networks, between simple and complex technologies. The information-theoretical argument underlying Structural Diversity nicely matches a primary motivation for looking at technological complexity from an economic perspective: since complex technologies entail more information, they are more difficult to learn, and to copy limiting their diffusion (Kogut and Zander, 1992). This makes these technologies more exclusive (or at least raises entry barriers), which in turn implies that actors that engage in complex technologies are more likely to extract (higher) rents from their application.

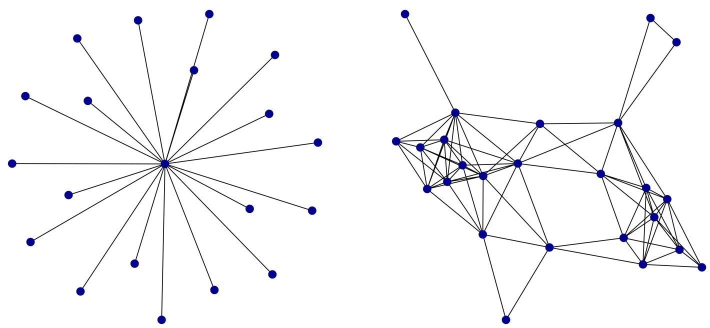
Unfortunately, there does not yet exist a direct quantification of the diversity of distinct topologies characterizing networks. However, the *Network Diversity Score (NDS)* measure developed by Emmert-Streib and Dehmer (2012) approximates this diversity in empirical settings. The NDS distinguishes between networks with ordered, complex, or with rather random structures. Ordered networks (e.g., a star) are characterized by few or even just one dominant topology. "Complex" network structures (e.g., a small-world network) involve multiple (and potentially overlaying) topologies. Lastly, at least theoretically, networks with near random structures are likely to have the greatest number of topologies[4]. Figure 1 provides visual impressions of a simple and a complex network structure.

Notably, the NDS measure has been developed to guarantee maximal differentiation between ordered, complex, and random networks by means of numerical optimization. It captures and quantifies these differences on a continuous scale, with small values indicating networks having high degrees of randomness, and larger values signaling the extent to which fewer distinct topologies shape these networks' structures. Crucially, it is this feature that resembles the idea of Structural Diversity and *not* the actual construction of the NDS index (Broekel, 2019).

In practice, the measure of Structural Diversity is calculated in multiple steps for each of the 655 4-digit technology classes of the *Cooperative Patent Classification* (CPC). Firstly, for each technology class c, we defined a set of nodes V, which represents all V 10-digit technology classes appearing

[3]This is a very simplified illustration to highlight the central idea and not an actually representation of a car.

[4]This is not to say that there are actually fully random combinatorial networks. Rather, networks with less-ordered structures tend to have higher levels of topological diversity.



(a) (b)

Figure 1: **(a)** A simple and **(b)** a complex network structure

on patents associated to technology c. Secondly, on the basis of these 10-digit classes' co-occurrences E on patents, the network $G_{c,e} = (V, E)$ is constructed for technology c. This combinatorial network $G_{c,e}$ entails all components that constitute technology c as well as all components that are combined with c (Broekel, 2019).

Next, the network is made binary with all positive links obtaining a value of one, and all non-existing links a value of zero. According to Emmert-Streib and Dehmer (2012), the network $G_{c,e}$ can be seen as one out of many realizations of an underlying (and unknown) network model G_c (i.e., it represents one individual network from a population of networks based on the same network model). It is this network model G_c that is actually of interest because it represents the mechanisms that create the (diversity of) topologies. Following Emmert-Streib and Dehmer (2012), the properties of G_c are captured by using the empirically observed network $G_{c,e}$ and calculating individual Network Diversity ($iNDS$) scores for a series of $G^S_{c,e}$ subnetworks extracted from $G_{c,e}$. In practice, we generated $G^S_{c,e}$ by using a random Walktrap algorithm with $w = 200$ steps based on a random sample of $S = 125$ nodes as a start[5]. The $iNDS$ for each subnetwork $G^S_{c,e}$ sampled from $G_{c,e}$ was calculated using equation 1, expressed as

$$iNDS(G^S_{c,e}) = \frac{\alpha_{module} * r_{graphlet}}{v_{module} * v_{\lambda}}, \tag{1}$$

where $\alpha_{module} = \frac{M}{n}$ represents the share of modules in the network, M is the number of modules, and V is the number of nodes. Modules are identified with the short random Walk approach by Pons and Latapy (2006). It is multiplied with the ratio of graphlets of sizes three and four, expressed as $r_{graphlet} = \frac{N_{graphlet}(3)}{N_{graphlet}(4)}$. The result is set into relation to the product of the variability of the network's Laplacian (L) matrix ($v_{\lambda} = \frac{var(\Lambda(L))}{mean(\Lambda(L))}$) and the variance of the module sizes m ($v_{module} = \frac{var(m)}{mean(m)}$) in the network.

The NDS as quantification of the diversity of topologies in G_c is obtained by averaging across the obtained $iNDS(G^S_{c,e})$, expressed as

$$NDS(\{G^S_{c,e}|G_{c,e}\}) = \frac{1}{S} \sum_{G^S_{c,e} \epsilon G_{c,e}} iNDS(G^S_{c,e}). \tag{2}$$

[5]In the case of a network with fewer than 125 nodes, S is equal to its number of nodes.

The measure of Structural Diversity cpx_c is calculated using equation 3, which is a simple transformation of the NDS, ensuring that large values correspond to higher levels of complexity and that the empirical values remain in an application-friendly range (Emmert-Streib and Dehmer, 2012):

$$cpx_c = log(\frac{1}{NDS(\{G^S_{c,e}|G_{c,e}\})}). \tag{3}$$

Note, in order to make the networks more stable, we used a three-year moving window approach, i.e., the combinatorial networks used to calculate cpx_c in year t are based on all patents in c in the years t to $t-2$. For each year, we obtained one individual complexity score for each of the 655 technologies (4-digit CPC classes). Table A.1 in Section A lists the most and least complex CPC classes in 2014.

3.1.2 Aggregating technological complexity at the regional level

Our central independent variable *regional complexity (rcpx)* represents the aggregation of technological complexity to the regional NUTS 2 level indicating regions' capabilities to create and utilize complex technologies. The literature does not provide a common approach of how to aggregate technology-specific measures (such as complexity) to the regional level. One straightforward approach is to calculate the average technological complexity of regional patents. However, the raw average might lead to sub-optimal outcomes. For instance, imagine a certain region that produces the most complex technologies. At the same time, a substantial share of its activities involve simple technologies. In this case, calculating the average complexity discriminates against complex technologies by considering simple ones. In fact, the information that this region is active in simple technologies does not provide any information about its capability to develop and manage complex technologies. We therefore refrain from using the raw average and calculate the average technological complexity in different percentiles of the regional complexity distribution.

Precisely, to calculate *rcpx*, we first assigned to each 4-digit CPC class c the corresponding complexity value cpx_c, as measured with Structural Diversity. Second, we assigned patents to European NUTS 2 regions using the residential information of inventors to avoid a potential bias of headquarters. Large corporations with various subsidiaries in multiple locations tend to file their patents through their headquarters even though the invention was developed in a subsidiary's location. Therefore, a "headquarters effect" can lead to systematic over- and underestimations of regional patenting activities. Using inventors' residences minimizes the statistical bias caused by headquarters. Third, for each region r and time t, we created an activity vector $A_{r,t}$ containing the set of CPC classes c that occur on patents of inventors from this region in that year. Crucially, if CPC classes appear on multiple patents (but not multiple times on the same patent), they are recorded as individual activities and are consequently kept as individual elements in the activity vector. Fourth, we sorted all CPC classes in the activity vector based on cpx_c in descending order, which yields the ranked complexity distribution of technological activities in regions beginning with the most complex CPC classes and ending with the least. Lastly, we calculated the aggregated regional complexity value $rcpx_{r,t}$ for each region by averaging across the subset of activities X ($X \subset A$), with X containing all complexity values $cpx_{c,r,t}$ ($c = 1, \ldots, n$) that belong to the percentile x of the regional complexity distribution, whereby n defines the size of the subset X (see equation 4):

$$rcpx_{r,t} = \frac{1}{n} \sum_{c \in X}^{n} cpx_{c,r,t} \tag{4}$$

For example, if x is 10, $rcpx_{r,t}$ is the average complexity of the top 10% most complex technological activities in region r and year t. Figure 2 visualizes our approach.

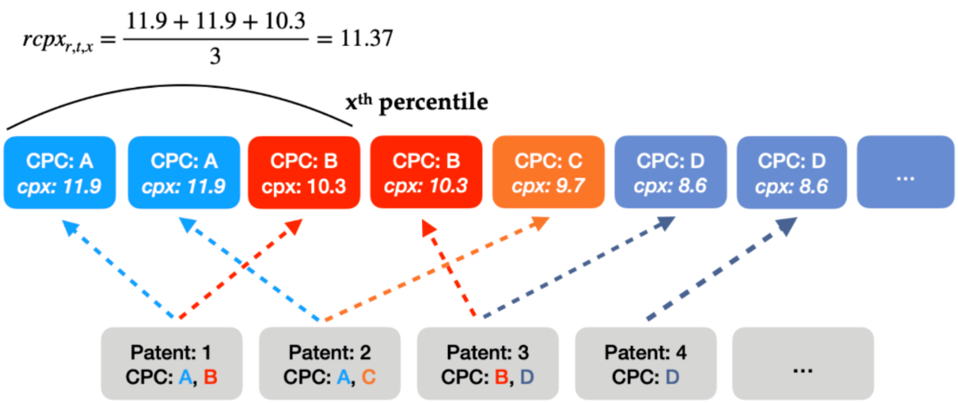


Figure 2: Aggregation of technological complexity at the regional level.

There is no obvious choice as to what percentile to choose in the empirical calculations. We therefore tested the robustness of our results with respect to a range of percentiles, from 1% to 25%.

3.2 Control variables

In addition to complexity, the literature has identified other determinants of regional growth and potential confounders of complexity, for which it is important to control. We distinguish between two sets of control variables: those that approximate regional technological capabilities, and those providing information on the local economic structure of regions.

Technological Capabilities

Innovation activities are an important source of regions' economic growth (Lucas, 1988; Romer, 1990). We consider the number of regional patents per capita to control for regional innovation capabilities. The long discussion about specialization and diversity indicates that not only size effects, but also the local technology structure, play a fundamental role in regional growth. This debate has not yet reached a final conclusion, and it rather seems that both specialization and diversity can be beneficial for the economic development of regions (Beaudry and Schiffauerova, 2009). We use the distribution of patents across technologies at the four-digit CPC level to indicate the degree of regional specialization and diversity, respectively. We follow the common approach in the literature and measure specialization as the average location quotient (Boschma et al., 2015). To approximate regional diversity, we relied on the Shannon entropy. The exponential of the individual entropy scores gives a diversity score, which is comparable across regions (Jost, 2006). Last, complexity is sometimes associated with high-tech activities (Eurostat, 2016). For example, Eurostat defines high-tech as a predetermined set of patent classes. To test complexity against this exogenous definition of high-tech activities, we included the regional share of patents in high-technologies (as defined by Eurostat (2016)) as an explanatory variable in the analysis.

Regional Economic Structure

We complemented our patent-based indicators with economic variables at the regional level, which were all collected from Eurostat. The literature on urban scaling has shown that populated places are more productive with respect to socio-economic outcomes such as GDP and innovation (Bettencourt et al., 2007). To control for these urbanization effects, we included population density as an explanatory variable. The availability of human capital in the form of highly educated people is also beneficial for regional growth (Lucas, 1988). Additionally, the increasing complexity of technologies requires better skilled labor. In line with previous studies, we used the share of people with a tertiary education as a proxy for human capital (Broekel, 2012). We also controlled for local unemployment rates, as higher rates are negatively associated with economic growth. Last, we included the share of employees in manufacturing, as patenting activities are biased towards manufacturing sectors. Table 1 summarizes all variables, their empirical definition, and their data sources. Basic descriptive statistics and correlations between these variables are reported in Table 2.

Table 1: List of variables with their definitions and data sources

Variable	**Definition**	**Data Source**
rcpx	Average regional complexity (in ln): [Average of the top % most complex activities (as measured with Structural Diversity) of the regional complexity distribution]	OECD REGPAT Database, own calculation
lgdp	Gross domestic product per inhabitant (in ln): [Total gross domestic product (purchasing power standards) divided by total population]	Eurostat
lpat	Total number of regional patents per capita (in ln): [Total number of patents / economically active working population aged 15 to 64 * 10,000]	OECD REGPAT Database, own calculation
lq	Average regional location quotient at the four-digit CPC level	OECD REGPAT Database, own calculation
div	Regional diversity measured as the exponential of the Shannon entropy of the regional patent distribution at the four-digit CPC level (Jost, 2006).	OECD REGPAT Database, own calculation
htec-pat	Share of patents in high-tech (in %): [Total regional patents in high-tech classes divided by total number of regional patents * 100]	OECD REGPAT Database, own calculation. Note: the classification is based on predefined technology classes considered as high-tech by Eurostat (2016)
lpopdens	Population density (in ln): [Economically active population aged 15-64 / Land area in square km]	Eurostat
hc	Human capital (in %) defined as: [Persons with tertiary education aged 25-64 divided by total population aged 25-64]	Eurostat
unemp	Regional unemployment rate (in %) defined as: [Unemployed persons divided by economically active population * 100]	Eurostat
share manufac	Employees in manufacturing (in %) defined as: [Employees in manufacturing divided by total number of employees * 100]	Eurostat

3.3 Estimation approach

We estimated a dynamic panel regression with region and time fixed effects to identify the relationship between regional complexity and economic growth in the following form:

$$lgdp_{r,t} = \beta_1 lgdp_{r,t-1} + \beta_2 rcpx_{r,t-l} + \gamma X_{r,t-l} + \phi_r + \upsilon_t + \mu_{r,t}. \quad (5)$$

Table 2: Descriptives and Correlations

	Variables	Min	Max	Mean	SD	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	lgdp	9.55	11.26	10.18	0.25									
2	rcpx	11.39	13.25	12.43	0.22	0.17***								
3	lpat	-0.95	4.01	2.24	0.82	0.52***	0.2***							
4	lq	1.31	46.06	4.56	4.09	-0.02	-0.36***	-0.22***						
5	div	26.25	264.32	129.23	42.25	0.22***	0.03	0.45***	-0.55***					
6	htec-pat	0.31	16	4.10	2.60	0.02	0.61***	0.04*	-0.23***	-0.19***				
7	lpopdens	0.38	7.67	4.45	1.13	0.26***	0.04	0.1***	-0.33***	0.21***	0.06**			
8	share-manufac	4.70	44.88	22.14	8.11	-0.22***	-0.2***	0.21***	0.01	0.26***	-0.2***	-0.15***		
9	hc	7	52.20	26.24	8.08	0.34***	0.28***	0.1***	-0.09***	-0.11***	0.3***	0.18***	-0.52***	
10	unemp	1.20	36.20	7.29	4.01	-0.35***	0.08***	-0.44***	-0.06**	-0.08***	0.06**	-0.02	-0.15***	0.03

Significance levels: *p<0.05; **p<0.01; ***p<0.001.

where the dependent variable ($lgdp_{r,t}$) is the regional GDP of region r in year t. Following the literature (Mohl and Hagen, 2010), the growth rate is expected to depend on the value of GDP per capita in the previous year $lgdp_{r,t-1}$. $rcpx_{r,t-l}$ represents the level of regional technological complexity in period $t-l$. $X_{r,t-l}$ is an $N \times K$ matrix of control variables. The corresponding $K \times 1$ vector γ contains the response parameters of our control variables. As mentioned above, we included regional and time fixed effects as denoted by ϕ_r and v_t, respectively, to account for unobserved time-invariant heterogeneity. $\mu_{r,t}$ denotes the error term.

When including a lagged version of the dependent variable as an explanatory variable in addition to region and time fixed effects, the OLS estimator suffers from the Nickell bias (Nickell, 1981). Frequently, an instrumental variable approach using GMM estimators are employed as a solution. However, applying GMM estimators requires strong assumptions about the appropriateness of past values of the dependent variable to function as valid instruments (Pickup et al., 2017). Moreover, these types of estimators are generally perceived as being relatively inefficient (Behr, 2003). Lancaster (2002) proposed an orthogonal reparameterization approach to obtain unbiased estimates from dynamic panel models, which is also known as the orthogonalized panel model (OPM). This approach has been shown to outperform the more common GMM estimators (Hsiao et al., 2002). In light of this, we employed the OPM regression as implemented by (Pickup et al., 2017) to estimate the parameters of our dynamic panel model.

4 Results

4.1 Complexity and regional growth

All results presented in this section are based on complexity being estimated with respect to the 10^{th} percentile. That is, a region's technological complexity corresponds to the mean Structural Diversity of its technological activities with the 10% highest values. Notably, our results are very robust with respect to the choice of this parameter (see Section 4.2). We also restricted our analysis to regions with at least 75 patents per year. This threshold is necessary to provide reliable results for all variables that are based on patent data[6]. Another important set of parameters to be specified are the time lags between GDP growth and the explanatory variables. For regional characteristics that are not based on patent data ($lpopdens$, $share\_manufac$, hc, $unemp$), we follow the literature and consider their values in year $t-1$ (Mohl and Hagen, 2010). In contrast, for variables that are based on patent information greater time lags, have been considered. First, patent data is known to represent innovation activities several years into the past (Acs et al., 2002). Second, it takes some time before innovations translate into economic growth. In the following, we present the results for a time lag of four years for the patent-based variables ($rcpx$, $lpat-pc$, lq, div, $htec-pat$). Alternative specifications of three and five years are discussed in Section 4.2.

Before we turn to our estimation outcomes, we present descriptive results regarding regional capabilities in complex technologies. Figure 3 shows the distribution of technological complexity across our sample of regions for the time period 2000-2014 (panels A and B). In panel A, values are grouped from low to high complexity using percentiles of the cross-regional complexity distribution. In general, high complexities are relatively scattered across the continent. Almost every country has at least one region in the highest complexity group, and, in many cases, this is the capital city or the region with the largest population. The south of Germany and large parts of Scandinavia, generally considered as highly R&D intensive with many technological leaders, represent agglomerations of highly complex technological capabilities.

[6]Due to the threshold of 75, 66 regions are removed from the sample. Our results, however, are not sensitive to the chosen patent threshold; see the robustness checks in Section 4.2

Investigating the regional distribution of technological complexity reveals interesting patterns of knowledge creation in regions that are hidden if only raw patent numbers are analyzed. For example, actors in Sicily (NUTS Code: ITG1) produced 1,078 patents in 2000-2014, which is below the European average of 3,059 patents (SD = 10,292, Median = 170). However, Sicily scores relatively high on technological complexity compared to other regions in Italy and Europe. The comparatively high complexity values are due to inventions related to semiconductors (CPC class H01L), which account for 14% of all patents in Sicily. With 12.34, the CPC class H01L has an above-average complexity score compared to the average of 10.56. The city of Catania, Sicily, hosts a nano-electronic cluster called *"Etna Valley,"* with *STMicroelectronics*, a leading company in the semiconductor industry, having R&D facilities in the region. The history of the cluster dates back to the 1960s, when STMicroelectronics decided to locate in Catania. As an anchor company, STMicroelectronics, together with a high-quality regional research system that includes universities (e.g., University of Messina and University of Palermo) and research institutions (e.g., *Institute of Microelectronics and Microsensors [IMM]*), attracted additional organizations and highly qualified employees (Baglieri et al., 2012), contributing to innovations in this complex technological field.

In sum, Sicily reaches high values in terms of technological complexity due to the presence of highly complex activities and consequently, it can be expected to show above-average economic growth. However, Sicily's economic growth between 2000 and 2014 (i.e., 0.8%) has been below the European average (i.e., 2.8%). This underlines the fact that technological complexity is one of many factors shaping regions' economic development. As the case of Sicily shows, depending on a region's specific situation, it may also not always be able to compensate for the lack of other growth determinants.

The panels C-E in Figure 3 visualize intra-regional complexity distributions in three selected regions in relation to the average across all regions. The city of Hamburg represents an urban area in Europe (panel C, *"Urban"*), East Anglia (including Oxfordshire) is a well-known R&D intensive region (panel D, *"Tech"*), and the economy of Agder og Rogaland in South-West Norway is specialized on extracting technologies in the oil and gas industry (panel E, *"Resource"*). Although Hamburg's complexity distribution shows a trend towards more complex activities, it also includes a wide spectrum of less complex activities representing the European average. East Anglia is characterized by a rather narrow range concentrated at the top end of the complexity distribution, supporting the region's image as an R&D hub. In contrast, the distribution of Agder og Rogaland has a wide range and is centered at relatively low complexity values compared with East Anglia and the European average. Accordingly, the regional complexity distribution illustrates structural differences in regions' technology profiles that correspond to their general technological and economic capabilities.

Figure 4 displays cross-regional dynamics of complexity over time. Panel A compares the complexity of regions in two consecutive time periods. Both values are highly correlated, as indicated by the correlation coefficient of 0.78, suggesting that regional complexity is relatively persistent over time. Panel B in Figure 4 further supports this observation by showing the correlation coefficient of consecutive annual complexity values, i.e., $rcpx_t$ and $rcpx_{t+1}$. Over a 15-year time period, the coefficient lies in the range of 0.90 and 0.96, indicating that regional complexity changed relatively slowly between 2000 and 2014. Nevertheless, regional technological complexity is not time-invariant; this is illustrated by panel A in Figure 4, which displays the top 10 regions with the highest positive and highest negative growth of complexity. Some regions managed to substantially increase their ability to produce complex technologies. For example, the German region Bremen (NUTS Code: DE50) experienced a 4% increase in complexity, although its patent output decreased in the same time period by 3%. This positive change primarily corresponds to increasing patenting activities in technologies with an above average technological complexity such as B64C (airoplanes, helicopters),

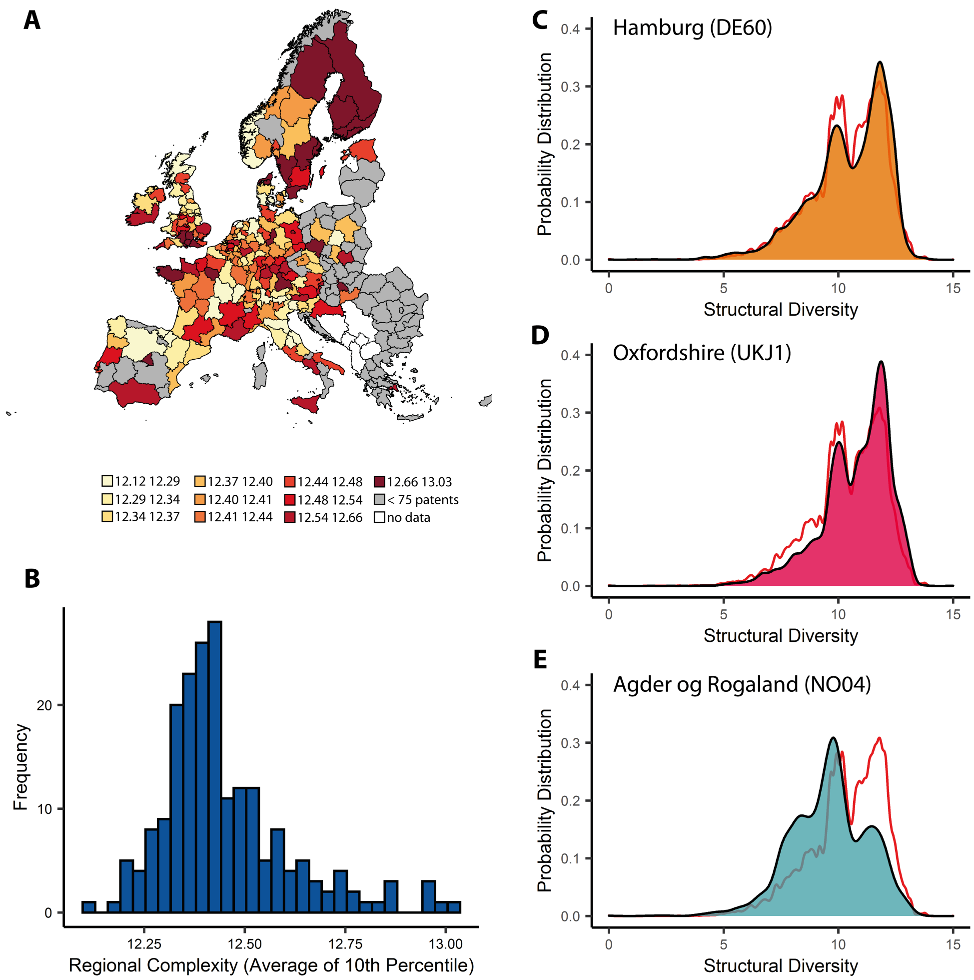


Figure 3: Technological complexity in Europe between 2000 and 2014. **A** Map and **B** distribution of regional complexity scores across all regions. Patents' complexity distribution in all regions (red line) compared with three selected regions **C** urban region (Hamburg, Germany), **D** technology-intensive region (Oxfordshire, United Kingdom), and **E** resource-intensive region (Agder og Rogaland, Norway).

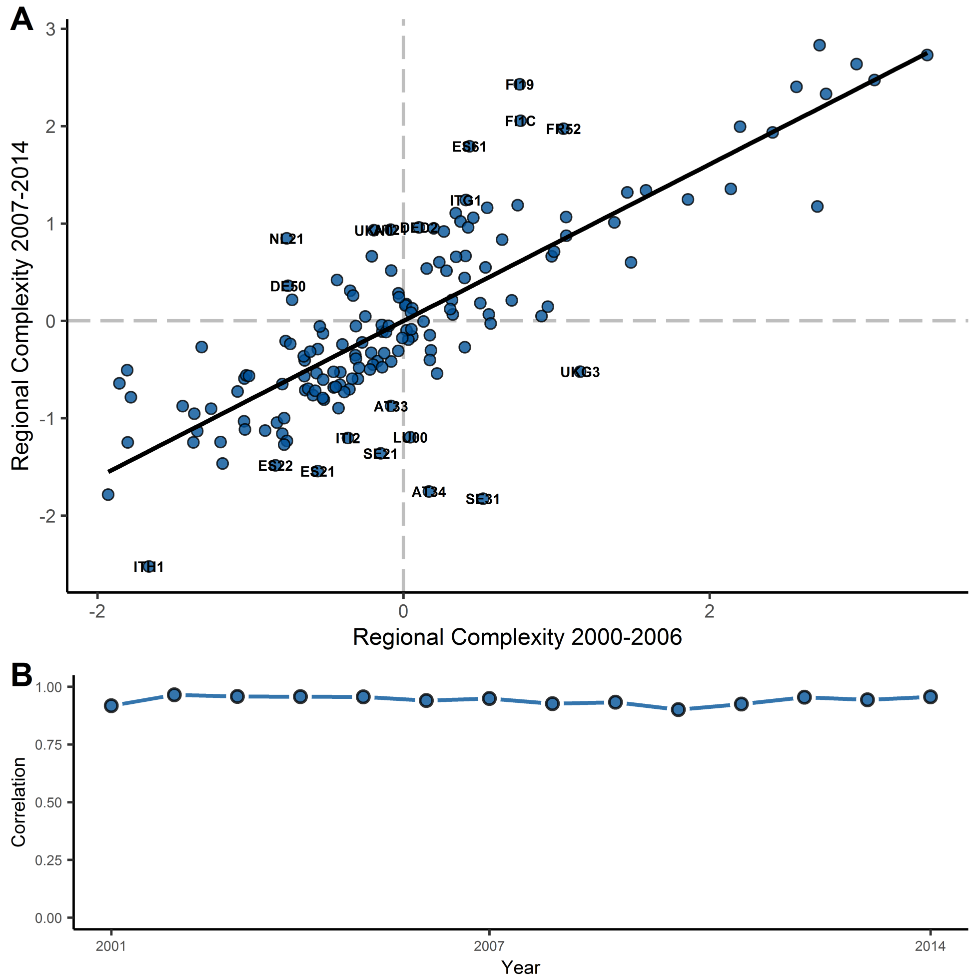


Figure 4: Cross-regional dynamics of complexity over time. **A** Regional complexity in two consecutive time periods. The corresponding correlation coefficient is 0.78. **B** Annual correlation of regional complexity between 2001 and 2014.

C08G (macromolecular compounds), and Y02T (climate change mitigation technologies).

To answer our main research question, we explain annual GDP growth with lagged values of regional technological complexity and additional covariates. Table 3 summarizes the results of our regression analyses. In our main estimation approach using the OPM dynamic panel regression, the coefficient of complexity ($rcpx$) is significantly positive. Accordingly, complexity is a positive and robust predictor of economic growth in NUTS 2 regions. The coefficient estimate takes values between 0.016 and 0.072, and is significant at the 99% level. This finding is supported by a conventional fixed effects regression using standard OLS.

Economic activities are not constrained by administrative boundaries such as NUTS 2 regions. Potential spatial spillovers between neighboring regions might lead to spatial autocorrelation in the errors. To address this, we applied spatial econometric regression techniques. The Lagrange Multiplier test suggested to use a spatial lag model with spatially correlated errors (Elhorst, 2014). In the definition of spatial weights, we followed existing approaches in the literature (Mohl and Hagen, 2010) and defined the weights matrix W based on a k-nearest neighbors approach, with

$k = 3$[7]. The results obtained with a GM estimator (Kapoor et al., 2007) are reported in column 3 of Table 3. The spatial lag variable ρ is significantly positive, suggesting that regions benefit from being located near regions with higher GDP per capita levels. The estimate of regional technological complexity *rcpx*, although smaller in magnitude, remains significantly positive at the 99% level.

Regional technological complexity and the dependent variable are measured on a logarithmic scale. Therefore, we can interpret the obtained coefficients as elasticities. Taking the coefficient of regional complexity in the OPM Model as a benchmark, a 1% increase in average regional complexity is associated with a 0.045% GDP per capita growth four years later. To put this into perspective, the average growth rate of complexity between 2000-2014 was 1.8%. Accordingly, a one percent increase in regional complexity represents a change that is close to the average growth of complexity over 15 years.

Table 3: Panel regression results for GDP growth

	OPM FE	OLS FE	Spatial FE
lgdp	0.843	0.696	0.667
	[0.712; 0.995]	[0.553; 0.839]	[0.618; 0.717]
rcpx	0.045	0.022	0.014
	[0.016; 0.072]	[0.012; 0.031]	[0.000; 0.027]
lpat	0.016	0.005	0.005
	[−0.015; 0.044]	[−0.017; 0.028]	[−0.007; 0.017]
htec-pat	−0.002	−0.001	−0.000
	[−0.005; 0.002]	[−0.003; 0.001]	[−0.002; 0.001]
lq	−0.000	−0.000	−0.001
	[−0.002; 0.002]	[−0.002; 0.001]	[−0.002; 0.000]
div	0.000	−0.000	−0.000
	[−0.000; 0.001]	[−0.000; 0.000]	[−0.000; 0.000]
lpopdens	−0.341	−0.038	−0.015
	[−0.480; −0.210]	[−0.129; 0.053]	[−0.075; 0.045]
share-manufac	0.003	0.000	0.000
	[0.002; 0.004]	[−0.002; 0.002]	[−0.001; 0.001]
hc	0.006	0.001	0.001
	[0.004; 0.008]	[−0.001; 0.003]	[−0.000; 0.002]
unemp	0.002	−0.004	−0.003
	[−0.001; 0.004]	[−0.006; −0.002]	[−0.004; −0.002]
ρ			0.099
			[0.002; 0.197]
Region FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	1,649	1,649	1,518

99% confidence intervals in parentheses.

OPM = Orthogonalized Panel Model.

Spatial panel model includes ρ as spatial autoregressive component.

[7] Alternative specifications with $k = 1$, $k = 5$, $k = 8$, and $k = 10$ yield similar results.

4.2 Robustness analysis

We conducted a number of additional analyses to evaluate the robustness of our findings with respect to three major parameters. First, we used the average technological complexity of the 10% most complex patents in a region to calculate regional technological complexity. Second, we restricted our sample to regions with more than 75 patents per year to increase the robustness of all patent-based variables ($rcpx$, $lpat$, lq, div, $htec - pat$). Third, we assumed a four-year time lag between GDP growth and the patent-based variables. To test the sensitivity of our findings with respect to these three parameters, we re-estimated the OPM regressions and varied all three parameters (complexity percentiles, patent thresholds, time lags) successively[8]. More precisely, we altered the aggregation of regional complexity and explored the averages of all percentiles between 1% and 25% of the most complex activities. We also varied the threshold for the minimum number of patents per year, testing all whole-integer values between 1 and 200. We repeated these to steps for three alternative time lag scenarios of three, four, and five years for all patent-based variables. Figure 5 visualizes the distribution of the lower bound of the 95% confidence intervals across these specifications. In the case of the three-year time lag (Panel A of Figure 5), complexity remains insignificant in all specifications. For the five-year time lag (Panel C of Figure 5), the range of significantly positive coefficients is reduced to specifications with complexity approximated by percentiles between one and ten percent. Nevertheless, our main results are confirmed for the time lag of five years as well. The four-year time lag (Panel B of Figure 5) is found to be the most robust, which motivated its use in the previous presentation of the results. A significantly positive coefficient of $rcpx$ is obtained for almost all patent-number thresholds up to 200. For larger values, the number of observations decreases too strongly to estimate reliable regressions. The significance of the coefficient is most pronounced within the range of two to fifteen percent percentiles. Notably, the coefficient does not become significantly negative in any of these specifications.

Last, we compared our results to those obtained with applied complexity measures. We also approximated regional technological complexity using (a) the Knowledge Complexity Indicator (KCI) as introduced by Balland and Rigby (2017) and used by Antonelli et al. (2020), which is an adaptation of the economic complexity index of Hidalgo and Hausmann (2009); and (b) the NK measure proposed and employed by Fleming and Sorenson (2001). The corresponding results are reported in Table 4 and paint an ambivalent picture. The point estimate of KCI is significantly positive at the 99% confidence level (first column), which supports our findings that complexity is beneficial for regional economic growth. However, the result is not robust when considering spatial dependencies (second column) or when using an alternative time lag specification.[9]. The estimated coefficients for NK complexity are insignificant in the OPM and in the spatial panel regression. Consequently, the choice of complexity measure matters in this context, which may explain the discrepancy between our results and those of Antonelli et al. (2020).

In summary, our dynamic panel regressions include a rich set of time-variant covariates and remain robust in many alternative specifications. They feature time and region fixed effects, capturing potential time-invariant variations. The findings also remain robust to the consideration of spatial dependencies and the inclusion of regional technological complexity in different time lags. We are therefore confident that our results support a causal interpretation of technological complexity being a driver of regional economic growth.

[8] In total, we calculated 1,000 (25 distinct percentiles and 40 alternative patent thresholds) regressions for each of the three time lags scenarios.

[9] When using a time lag of five years, the coefficient of KCI becomes significantly negative. The results can be obtained from the authors upon request.

Table 4: Robustness analysis using KCI and NK as alternative complexity measures

	KCI OPM	KCI Spatial	NK OPM	NK Spatial
lgdp	0.987	0.969	0.672	0.671
	[0.916; 0.995]	[0.622; 0.722]	[0.753; 0.995]	[0.622; 0.721]
kci	0.001	0.001		
	[0.001; 0.001]	[−0.001; 0.001]		
nk			0.000	−0.003
			[−0.015; 0.016]	[−0.010; 0.003]
lpat_pc	−0.004	0.006	0.004	0.005
	[−0.027; 0.020]	[−0.008; 0.016]	[−0.021; 0.036]	[−0.007; 0.017]
htec_pat	0.001	0.001	0.000	0.000
	[−0.002; 0.004]	[−0.001; 0.002]	[−0.002; 0.004]	[−0.001; 0.002]
lq	−0.001	−0.001	−0.001	−0.001
	[−0.003; 0.001]	[−0.002; 0.000]	[−0.003; 0.001]	[−0.002; 0.000]
div	0.000	0.000	−0.000	−0.000
	[−0.000; 0.000]	[−0.000; 0.000]	[−0.000; 0.001]	[−0.000; 0.000]
lpopdens	−0.358	−0.367	−0.018	−0.018
	[−0.475; −0.243]	[−0.078; 0.042]	[−0.496; −0.239]	[−0.078; 0.041]
share_manufac	0.001	0.003	0.000	0.000
	[0.000; 0.002]	[−0.001; 0.001]	[0.002; 0.004]	[−0.001; 0.001]
hc	0.005	0.005	0.001	0.001
	[0.003; 0.006]	[−0.000; 0.002]	[0.004; 0.007]	[−0.000; 0.002]
unemp	0.002	0.003	−0.003	−0.003
	[0.001; 0.003]	[−0.004; −0.002]	[−0.000; 0.004]	[−0.004; −0.002]
ρ		0.098		0.095
		[−0.000; 0.196]		[−0.004; 0.193]
Region FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	1,649	1,518	1,649	1,518

99% confidence intervals in parentheses.

OPM = Orthogonalized Panel Model.

Spatial panel model includes ρ as spatial autoregressive component.

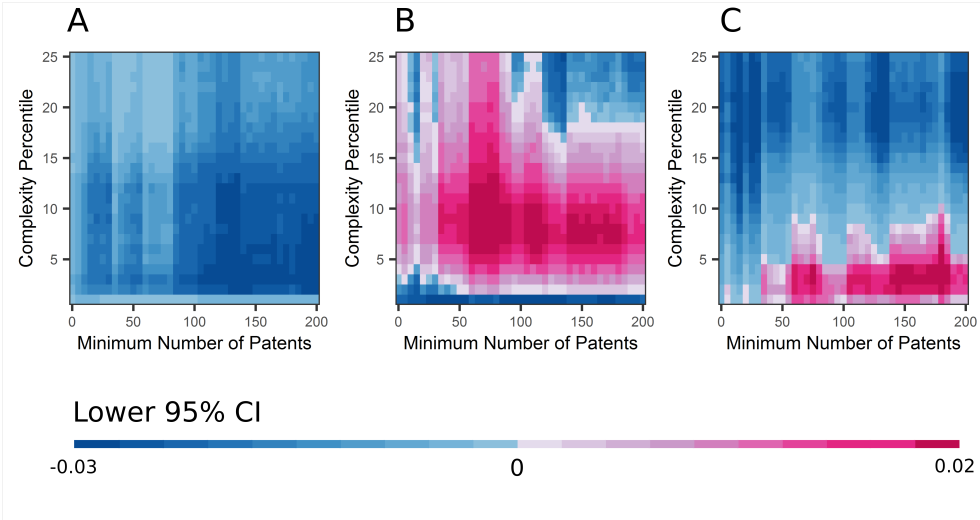


Figure 5: Robustness analysis with alternative complexity percentiles (ranging between top 1% and top 25% complex activities) and minimum patent numbers (ranging between 0 and 200 patents). Lower 95% confidence intervals are based on OPM estimations. Pink colors indicate that the lower boundary of the confidence intervals excludes values equal to and smaller than zero. $rcpx$ is lagged by **A** 3 years, **B** 4 years and **C** 5 years.

5 Conclusion

In this article, we analyzed the contribution of technological complexity to regions' economic growth. Using data on European regions and a range of empirical specifications, we showed that differences in the capability to produce and exploit complex technologies explain variations in regions' economic growth. These findings underpin the argument that complexity is an important building block of competitive advantage (Kogut and Zander, 1992; Hidalgo and Hausmann, 2009).

However, there are a number of limitations that need to be taken into consideration when interpreting the results. Due to data limitations, our analysis was restricted to a specific time horizon of 15 years. Considering the longevity of economic development, 15 years might be too short a time frame to capture all aspects of the relationship between economic development and technological complexity. It is likely that technological complexity unfolds its effects on economic growth over even longer time periods (Fink et al., 2017), or that both economic development and the evolution of complexity interact in a co-evolutionary process spanning decades.

Another limitation is that our analysis was based on NUTS 2 regions in Europe. Although NUTS 2 regions are important entities for regional policy decisions, they represent administrative rather than functional spatial units. Functional regions in terms of metropolitan areas or labor market regions are often used in empirical analyses to limit spatial biases, for instance, due to commuting patterns. Future research should replicate our study using functional regions to ensure the robustness of our findings for different spatial units and scales.

Although our robustness checks underline the importance of technological complexity for regional economic growth, they are restricted to technological complexity being measured with Structural Diversity. Using two alternative complexity indicators yielded ambivalent results, highlighting a crucial challenge in the contemporary literature. Existing empirical investigations employ a variety of complexity measures, impeding the comparison of empirical results across studies. While the ECI/KCI approach has recently gained popularity, previous works have criticized its methodological basis: the method of reflection (Tacchella et al., 2012). As underlined by our study as well as by the works of Antonelli et al. (2020) and Broekel (2019), the ECI/KCI does not seem to be straightforwardly applied to European (regional) patent data. Clearly, more methodological research is necessary to improve our understanding about how to empirically capture and quantify technological complexity.

Nevertheless, our results fuel a number of important discussions. By providing empirical evidence of technological complexity impacting regional economic growth, our findings support the idea of building competitive advantages in complex activities. Technological complexity grows over time, demanding higher qualified individuals and more intensive collaboration (Powell et al., 1996; Pintea and Thompson, 2007; Wuchty et al., 2007; Broekel, 2019). Consequently, places that attract qualified individuals and that are embedded in interregional knowledge networks are better positioned to follow this strategy. As put forward by Balland et al. (2020), this is likely to amplify the geographic concentration of complex innovation activities even more, and it might be one of the reasons why urban agglomerations are increasingly becoming the centers of innovation. Our results add to this and confirm some of the empirical evidence of Balland et al. (2020). Specifically, our study confirms that complex technologies have a tendency to concentrate in large metropolitan areas (e.g., Paris, Madrid, Berlin, Stockholm, Munich). Complex knowledge requires, on average, more complementary factors than simpler knowledge. Large cities provide access to such factors at a relatively small spatial scale, which explains the spatial concentration of complex knowledge in urban areas (Gomez-Lievano et al., 2017).

However, we also showed that complexity is not restricted to urban agglomerations. There are many non-metropolitan regions that are able to develop complex technologies. At this point, it is unclear if this discrepancy is due to differences between the USA, on which the study of Balland et al. (2020) is based, and Europe, which is the focus of the present study. Alternatively, it can also be related to the application of different measures of complexity in the two studies. Clearly, this calls for more research investigating the evolution and drivers of complexity at the level of technologies and regions in future research.

Despite these unresolved (empirical) issues, technological complexity has already entered contemporary policy debates (Balland et al., 2019). In this vein, the complexity of sectors and technologies represents an ambivalent concept for policy makers. Nowadays, policy requires regions to invest in promising diversification strategies as evident in the smart specialization strategy of the EU facilitating regional development (Foray et al., 2011). Balland et al. (2019) argue that the combination of complexity and relatedness provides a promising concept to derive such smart diversification strategies. Accordingly, building a regional competitive advantage into new activities is argued to be beneficial (i.e., complex activities) and feasible (i.e., related activities) for regions. Besides its positive impact on developing technological strengths (Balland et al., 2019), our study provides empirical evidence that such a policy is likely to directly facilitate economic growth.

Yet, it is still unclear how regions can exactly build competitive advantage in complex activities and if this strategy is suited to, and desirable for every region. Of similar relevance is the question as to whether - and if so, how - policy can support the upgrading of regional capabilities to higher levels of complexity. As the increasing complexity of knowledge production demands better qualified individuals and more collaboration, programs targeting these are promising candidates in this regard. For instance, the EU *Framework Programme* (FP) appears to be a good tool in this context, as it is explicitly designed to facilitate knowledge and expertise exchange between regions. Their monetary incentives may help in overcoming barriers of knowledge diffusion that are particularly pronounced in the case of complex knowledge (Balland and Rigby, 2017). However, so far, evaluations of innovation policies largely neglect the dimension of technological complexity. The results of our study demonstrate that such neglect implies ignoring an important determinant of regional economic growth.

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):1069–1085.

Aghion, P. and Howitt, P. (1998). *Endogenous growth theory*. MIT Press, Cambridge, Mass.

Antonelli, C., Crespi, F., and Quatraro, F. (2020). Knowledge complexity and the mechanisms of knowledge generation and exploitation: The European evidence. *Research Policy*, page 104081.

Arthur, W. B. (2009). *The nature of technology: what it is and how it evolves*. Free Press, New York. OCLC: 6ocm14315023.

Aunger, R. (2010). Types of technology. *Technological Forecasting and Social Change*, 77(5):762–782.

Baglieri, D., Cinici, M. C., and Mangematin, V. (2012). Rejuvenating clusters with 'sleeping anchors': The case of nanoclusters. *Technovation*, 32(3-4):245–256.

Bahar, D., Hausmann, R., and Hidalgo, C. A. (2014). Neighbors and the evolution of the comparative advantage of nations: Evidence of international knowledge diffusion? *Journal of International Economics*, 92(1):111–123.

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

Balland, P.-A., Jara-Figueroa, C., Petralia, S. G., Steijn, M. P. A., Rigby, D. L., and Hidalgo, C. A. (2020). Complex economic activities concentrate in large cities. *Nature Human Behaviour*, 4(3):248–254.

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

Beaudry, C. and Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Research Policy*, 38(2):318–337.

Becker, G. S., Glaeser, E. L., and Murphy, K. M. (1999). Population and Economic Growth. *American Economic Review*, 89(2):145–149.

Becker, G. S. and Murphy, K. M. (1992). The Division of Labor, Coordination Costs, and Knowledge. *The Quarterly Journal of Economics*, 107(4):1137–1160.

Behr, A. (2003). A Comparison of Dynamic Panel Data Estimators: Monte Carlo Evidence and an Application to the Investment Function. *Bundesbank Discussion Paper*, 03(05):1–33.

Bettencourt, L. M., Lobo, J., and Strumsky, D. (2007). Invention in the city: Increasing returns to patenting as a scaling function of metropolitan size. *Research Policy*, 36(1):107–120.

Boschma, R. (2005). Proximity and Innovation: A Critical Assessment. *Regional Studies*, 39(1):61–74.

Boschma, R., Balland, P.-A., and 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. and Lissoni, F. (2009). Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows. *Journal of Economic Geography*, 9(4):439–468.

Breschi, S. and Malerba, F. (1997). Sectoral Innovation Systems: Technological Regimes, Schumpeterian Dynamic, and Spatial Boundaries. In *Systems of Innovation*, pages 130–156. Routledge, London, New York.

Broekel, T. (2012). Collaboration Intensity and Regional Innovation Efficiency in Germany—A Conditional Efficiency Approach. *Industry & Innovation*, 19(2):155–179.

Broekel, T. (2019). Using structural diversity to measure the complexity of technologies. *PLOS ONE*, 14(5):e0216856.

Elhorst, J. P. (2014). Spatial Panel Data Models. In *Spatial Econometrics*, pages 37–93. Springer Berlin Heidelberg, Berlin, Heidelberg. Series Title: SpringerBriefs in Regional Science.

Emmert-Streib, F. and Dehmer, M. (2012). Exploring Statistical and Population Aspects of Network Complexity. *PLoS ONE*, 7(5):e34523.

Eurostat (2016). High-tech industry and knowledge-intensive services (htec).

Fagerberg, J., Verspagen, B., and Caniëls, M. (1997). Technology, Growth and Unemployment across European Regions. *Regional Studies*, 31(5):457–466.

Feldman, M. P. (1994). *The Geography of Innovation*, volume 2 of *Economics of Science, Technology and Innovation*. Springer Netherlands, Dordrecht.

Ferrarini, B. and Scaramozzino, P. (2016). Production complexity, adaptability and economic growth. *Structural Change and Economic Dynamics*, 37:52–61.

Fink, T. M. A., Reeves, M., Palma, R., and Farr, R. S. (2017). Serendipity and strategy in rapid innovation. *Nature Communications*, 8(1).

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

Fleming, L. and Sorenson, O. (2004). Science as a map in technological search. *Strategic Management Journal*, 25(8-9):909–928.

Foray, D., David, P. A., and 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. Technical Report 001, Management of Technology and Entrepreneurship Institute.

Gertler, M. S. (2003). Tacit knowledge and the economic geography of context, or The undefinable tacitness of being (there). *Journal of Economic Geography*, 3(1):75–99.

Glaeser, E. L., Kallal, H. D., and Scheinkman, J. A. (1992). Growth in Cities. *Journal of Political Economy*, 100(6):1126–1152.

Gomez-Lievano, A., Patterson-Lomba, O., and Hausmann, R. (2017). Explaining the prevalence, scaling and variance of urban phenomena. *Nature Human Behaviour*, 1(1).

Griliches, Z. (1990). Patent Statistics as Economic Indicators: A Survey. Technical Report 3301, National Bureau of Economic Research, Cambridge, MA.

Grossman, G. M. and Helpman, E. (1991). Quality Ladders in the Theory of Growth. *The Review of Economic Studies*, 58(1):43.

Hargadon, A. (2003). *How breakthroughs happen: the surprising truth about how companies innovate.* Harvard Business School Press, Boston, Mass.

Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., Chung, S., Jiminez, J., Simoes, A., and Yildirim, M. A. (2013). *The Atlas of economic complexity: mapping paths to prosperity.* Center for International Development, Harvard University, Cambridge. OCLC: 961922687.

Henderson, J. V., Shalizi, Z., and Venables, A. J. (2001). Geography and development. *Journal of Economic Geography*, 1(1):81–105.

Henrich, J. (2004). Demography and Cultural Evolution: How Adaptive Cultural Processes Can Produce Maladaptive Losses - The Tasmanian Case. *American Antiquity*, 69(02):197–214.

Hidalgo, C. A. and Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*, 106(26):10570–10575.

Hsiao, C., Hashem Pesaran, M., and Kamil Tahmiscioglu, A. (2002). Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods. *Journal of Econometrics*, 109(1):107–150.

Jaffe, A. B., Trajtenberg, M., and Henderson, R. (1993). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *The Quarterly Journal of Economics*, 108(3):577–598.

Jost, L. (2006). Entropy and Diversity. *Oikos*, 113(2):363–375.

Kapoor, M., Kelejian, H. H., and Prucha, I. R. (2007). Panel data models with spatially correlated error components. *Journal of Econometrics*, 140(1):97–130.

Kauffman, S. A. (1993). *The Origins of Order - Self-Organization and Selection in Evolution.* Oxford University Press, New York.

Kogut, B. and Zander, U. (1992). Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. *Organization Science*, 3(3):383–397.

Kremer, M. (1993). Population Growth and Technological Change: One Million B.C. to 1990. *The Quarterly Journal of Economics*, 108(3):681–716.

Kuznets, S. (1962). Inventive Activity: Problems of Definition and Measurement. In National Bureau of Economic Research, editor, *The Rate and Direction of Inventive Activity: Economic and Social Factors*, pages 19–52. Princeton University Press.

Lancaster, T. (2002). Orthogonal Parameters and Panel Data. *Review of Economic Studies*, 69(3):647–666.

Lawson, C. and Lorenz, E. (1999). Collective Learning, Tacit Knowledge and Regional Innovative Capacity. *Regional Studies*, 33(4):305–317.

London, J. (1904). *The Sea Wolf.* Macmillan.

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22(1):3–42.

Malmberg, A. and Power, D. (2005). (How) Do (Firms in) Clusters Create Knowledge? *Industry & Innovation*, 12(4):409–431.

Markusen, A. (1996). Sticky Places in Slippery Space: A Typology of Industrial Districts. *Economic Geography*, 72(3):293.

McEvily, S. K. and Chakravarthy, B. (2002). The persistence of knowledge-based advantage: an empirical test for product performance and technological knowledge. *Strategic Management Journal*, 23(4):285–305.

Mohl, P. and Hagen, T. (2010). Do EU structural funds promote regional growth? New evidence from various panel data approaches. *Regional Science and Urban Economics*, 40(5):353–365.

Nelson, R. R. and Winter, S. G. (1982). *An evolutionary theory of economic change*. The Belknap Press of Harvard Univ. Press, Cambridge, Mass. OCLC: 255191816.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica*, 49(6):1417.

Petralia, S., Balland, P.-A., and Morrison, A. (2017). Climbing the ladder of technological development. *Research Policy*, 46(5):956–969.

Pickup, M., Gustafson, P., Cubranic, D., and Evans, G. (2017). OrthoPanels: An R Package for Estimating a Dynamic Panel Model with Fixed Effects Using the Orthogonal Reparameterization Approach. *The R Journal*, 9(1):60.

Pintea, M. and Thompson, P. (2007). Technological complexity and economic growth. *Review of Economic Dynamics*, 10(2):276–293.

Polanyi, M. (1966). *The Tacit Dimension*. Anchor Day Books, New York.

Pons, P. and Latapy, M. (2006). Computing communities in large networks using random walks. *Journal of Graph Algorithms and Applications*, 10(2):191–218.

Powell, W. W., Koput, K. W., and Smith-Doerr, L. (1996). Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology. *Administrative Science Quarterly*, 41(1):116.

Rivkin, J. W. (2000). Imitation of Complex Strategies. *Management Science*, 46(6):824–844.

Romer, P. M. (1990). Endogenous Technological Change. *Journal of Political Economy*, 98(5, Part 2):S71–S102.

Simon, H. A. (1962). The Architecture of Complexity. *Proceedings of the American Philosophical Society*, 106(6):467–482.

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. W. Strahan and T. Cadell, London.

Sorenson, O., Rivkin, J. W., and Fleming, L. (2006). Complexity, networks and knowledge flow. *Research Policy*, 35(7):994–1017.

Stojkoski, V., Utkovski, Z., and Kocarev, L. (2016). The Impact of Services on Economic Complexity: Service Sophistication as Route for Economic Growth. *PLOS ONE*, 11(8):1–29.

Storper, M. (2010). Why Does a City Grow? Specialisation, Human Capital or Institutions? *Urban Studies*, 47(10):2027–2050.

Sveikauskas, L. (1975). The Productivity of Cities. *The Quarterly Journal of Economics*, 89(3):393.

Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A., and Pietronero, L. (2012). A New Metrics for Countries' Fitness and Products' Complexity. *Scientific Reports*, 2(1).

Teece, D. J. (1977). Technology Transfer by Multinational Corporations: The Resource Cost of Transferring Technological Know-How. *The Economic Journal*, 87(346):242–261.

Teece, D. J., Pisano, G., and Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, 18(7):509–533.

Usher, A. P. (1954). *A history of mechanical inventions*. Dover, New York.

Winter, S. G. (1987). Knowledge and Competence as Strategic Assets. In *The Competitive Challenge: Strategies for Industrial Innovation and Renewal*, pages 165–187. Harper & Row, Ballinger Division, New York.

Wuchty, S., Jones, B. F., and Uzzi, B. (2007). The Increasing Dominance of Teams in Production of Knowledge. *Science*, 316(5827):1036–1039.

Yayavaram, S. and Chen, W.-R. (2015). Changes in firm knowledge couplings and firm innovation performance: The moderating role of technological complexity: Changes in Knowledge Couplings and Innovation Performance. *Strategic Management Journal*, 36(3):377–396.

Zander, U. and Kogut, B. (1995). Knowledge and the Speed of the Transfer and Imitation of Organizational Capabilities: An Empirical Test. *Organization Science*, 6(1):76–92.

A Most and least complex CPC classes

Table A.1: Ranking of the most and least complex CPC classes 2014

Rank	**CPC Class**	**Description**
1	B60L	propulsion of electrically-propelled vehicles
2	H04W	wireless communication networks
3	Y04S	systems integrating technologies related to power network operation, communication or information technologies for improving the electrical power generation, transmission, distribution, management or usage, i.e. smart grids
4	B33Y	additive manufacturing, i.e. manufacturing of three-dimensional [3-d] objects by additive deposition, additive agglomeration or additive layering, e.g. by 3-d printing, stereolithography or selective laser sintering
5	A61H	physical therapy apparatus, e.g. devices for locating or stimulating reflex points in the body; artificial respiration; massage; bathing devices for special therapeutic or hygienic purposes or specific parts of the body
6	B60W	conjoint control of vehicle sub-units of different type or different function; control systems specially adapted for hybrid vehicles; road vehicle drive control systems for purposes not related to the control of a particular sub-unit
7	F05D	indexing scheme for aspects relating to non-positive-displacement machines or engines, gas-turbines or jet-propulsion plants
8	F01D	non-positive displacement machines or engines, e.g. steam turbines
9	H03F	amplifiers
10	C10N	indexing scheme to lubricating compositions
⋮	⋮	⋮
614	F22G	superheating of steam
615	F15C	fluid-circuit elements predominantly used for computing or control purposes
616	G10F	automatic musical instruments
617	B27F	dovetailed work; tenons; slotting machines for wood or similar material; nailing or stapling machines
618	D02H	warping, beaming or leasing
619	A42C	manufacturing or trimming hats or other head coverings
620	B68B	harness; devices used in connection therewith; whips or the like
621	C12J	vinegar; its preparation
622	B68C	saddles; stirrups
623	B61J	shifting or shunting of rail vehicles