Relatedness, Cross-relatedness and Regional Innovation Specializations: An Analysis of Technology, Design and Market Activities in Europe and the US

Carolina Castaldi & Kyriakos Drivas

Papers in Evolutionary Economic Geography

23.07

Utrecht University Human Geography and Planning

Relatedness, Cross-relatedness and Regional Innovation Specializations: An Analysis of Technology, Design and Market Activities in Europe and the US

Carolina Castaldi
Department of Human Geography and Spatial Planning
Utrecht University
c.castaldi@uu.nl

Kyriakos Drivas
Department of Economics
University of Piraeus
drivas@unipi.gr

Abstract

This paper examines how regions develop new innovation specializations, covering different activities in the whole process from technological invention to commercialization. We develop a conceptual framework anchored in two building blocks: first, the conceptualization of innovation as a process spanning technology, design and market activities; second, the application and extension of the principle of relatedness to understand developments within and between the different innovation activities. We offer an empirical investigation where we operationalize the different innovation activities using three intellectual property rights (IPRs): patents, industrial designs and trademarks. We provide two separate analyses of how relatedness and cross-relatedness matter for the emergence of new specializations: for 259 NUTS-2 European regions and for 363 MSAs of the US. While relatedness is significantly associated with new regional specializations for all three innovation activities, cross-relatedness between activities also plays a significant role. Our study has important policy implications for developing and monitoring Smart Specialization regional strategies.

JEL codes*:* O34, O38, R11

Keywords: innovation, relatedness, regional specialization, patents, trademarks, designs, NUTS-2 regions, Metropolitan Statistical Areas.

Note: this is a working paper version of a forthcoming paper, please cite the published version: Castaldi, C. and Drivas, K. (2023), Relatedness, Cross-relatedness and Regional Innovation Specializations: An Analysis of Technology, Design and Market Activities in Europe and the US, *Economic Geography*, https://doi.org/10.1080/00130095.2023.2187374

Introduction

How innovation unfolds in space and over time is a critical question in understanding the ways in which regions can reconfigure their activities and thrive (Feldman, 1994). However, what is innovation? Schumpeter (1934) already stressed that invention is not yet innovation: much needs to happen before a novel idea turns into an actual new product or process that can generate value for users and producers alike. Nevertheless, most conceptual and empirical research on the geography of innovation has examined upstream and downstream stages of the innovation process in isolation, with a predominant focus on the former; i.e. technological invention. Turning invention into innovation also requires capabilities such as design and marketing (Mendonça, 2014; Rodríguez-Pose and Lee, 2020), which are critical for developing a persuasive innovation that is more likely to be adopted. Hence focusing on technology alone can result in a misrepresentation of the innovation process, leading to a bias in the preferred policy options too (Breznitz, 2021). Regional innovation systems can combine capabilities in all the complementary activities needed for innovation, but they also often specialize in specific ones and not all regions can or wish to be technology leaders (Asheim and Coenen, 2005; Capello and Lenzi, 2013).

Recognizing the different specializations open to regions is also at the core of policies towards Smart Specialization (Foray et al. 2011, 2014). The program aims at taking a broad view on innovation. Yet, regional policymakers and scholars alike struggle to capture the diversity of innovation specializations and end up focusing on one type of specialization at a time (Foray et al., 2018). Most often the focus is on specialization in science and technology activities: this can be explained with the belief that investing in the upstream stages of innovation will naturally lead to all kinds of innovations being introduced in the market (Marques and Morgan, 2018), but also with the fact that those innovation activities have been easier to monitor with data (Castaldi and Mendonça, 2022). All this has left many regions, especially those that are not high-tech clusters, struggling with recognizing and valuing their specific innovation capabilities and their ability to build smart specializations from them (Radosevic, 2018).

The objective of this paper is to develop a conceptual framework to understand the emergence of regional innovation specializations spanning a broader set of innovation activities than those focused on technological invention only. Our aim is to offer a framework that resonates with insights from theorizing and empirical findings of prior research, while being applicable in quantitative analyses of regional specializations, for policy and research purposes alike.

For the theoretical embedding, we leverage and extend the principle of relatedness (Hidalgo et al. 2018) and insights from evolutionary economic geography on related diversification (Boschma, 2017). To do so, a first conceptual step relies on clarifying the distinction between invention and innovation: we propose to separate three different activities, namely technology, design and market ones. For each activity we discuss the key properties and conceptualize the underlying knowledge space, to then discuss how relatedness can be defined in each space. A second conceptual step involves connecting the three innovation activities by introducing the idea of 'cross-relatedness' and the possibility to capture an overall innovation space. By 'space' we refer to a network where one can represent which innovation activities tend to co-specialize at the regional level. The co-specialization is depicted as a connection, with the innovation activities being the nodes in the network. Patterns of relatedness ('within' each innovation activity) and cross-relatedness ('between' innovation activity) can then be used to model the emergence of new regional innovation specializations, of the three different kinds.

For the operationalization, we propose a comparable set of innovation metrics that have not been systematically combined in regional innovation studies before. We capture the three innovation activities by three types of intellectual property rights (IPRs): i) (utility) patents, ii) industrial designs and iii) trademarks. These data allow to operationalize relatedness and cross-relatedness using the underlying patent, design and trademark classifications. We apply our empirical model to two independent settings for the period 2003-2016. The first is 259 NUTS-2 regions across 21 European countries: this setting is the most salient in relation to Smart Specialization policy applications. The second is 363 Metropolitan Statistical Areas (MSAs) areas in the United States (US), an alternative testbed where the different definition of industrial designs enables us to compare the role of two specific types of design activities, namely technical and aesthetic ones.

We reveal three main patterns. First, relatedness plays a significant role in the emergence of new regional specializations for all three innovation activities, not only the upstream ones. Second, cross-relatedness of technology activities with downstream ones matters for the emergence of design and market specialization, in line with traditional technology-push models of innovation. Yet, cross-relatedness of design and market activities with technology also matters for the emergence of new specialization, albeit to a lesser extent than relatedness in the same innovation activity. Nonetheless these 'backward linkages' are indicative of feedback loops and synergies between regional innovation activities. Finally, the comparison of the European and US contexts highlights the differential role of technical vs aesthetic design activities, while also informing the use of IPR metrics to capture such activities.

Regional innovation and the principle of relatedness: towards a conceptual framework

Unpacking innovation: technology, design and market activities

Innovation is more than invention: it requires turning a promising new idea into something that users are willing to buy or adopt. We conceptualize this process as made of three main activities: technology, design and market ones.

New technology typically stems from dedicated research activities, which can be formal R&D or informal on the job activities. The knowledge involved is often synthetic, typical of engineering sciences (Asheim and Coenen*,* 2005). Engineers and other technology developers will consider options across a 'technology space' (Dosi, 1997). A rich empirical literature has used patent data to reconstruct how companies and/or regions navigate the underlying technology space by following clear trajectories of learning (e.g. Leten et al. 2007; Rigby, 2015), showing a high degree of path-dependence.

Eventually, the new technological options can lead to a new product or a new process, but these will have to be further developed and designed before they can actually be applied and used. Design activities include prototyping and try-outs. When presented with new technological options, designers will work by navigating alternative design options in what can be defined as a 'design space' (Windrum et al. 2017). Design can be seen as an intermediary function, concerned with finding solutions to trade-offs between technical feasibility and users' preferences (d'Ippolito, 2014). Such a role is highly specific to technology-driven innovation processes, where designers are typically called upon only once new technologies emerge. Instead, design can take a more leading role in innovation processes typical of industries where soft innovation is the main source of change (Stoneman, 2010). There, designers focus on aesthetic design options and the creation of new meanings often with the aim of initiating new product lines and allow differentiation (Verganti, 2008).

Working product or process configurations will find their way to the market in the commercialization stage. In this last stage of the innovation process, capabilities related to marketing appear crucial. The success of an innovation depends not only on the quality of the innovation but also upon the extent to which it aligns with needs and aspirations of consumers. In this phase, symbolic knowledge related to the definition of new categories and meanings comes into play (Mendonça et al. 2004). When positioning the innovation, firms will consider profiling their offering as a specific option in a what can be called the 'market space'. To illustrate the point, Davids and Frenken (2018) reconstruct how Unilever positioned the margarine as a (healthy) food product after being first introduced as a medical product. Patterns of regional and corporate market diversification and specialization can be captured with trademarks (Castaldi and Mendonça, 2022).

Table 1 summarizes the key features of the three innovation activities, as the building blocks for our framework of regional innovation specialization. Based on the discussion above, design activities can be of two kinds. A first kind concerns what one could call 'technical design': these activities are common in technology-driven innovation processes and often involve designers trained at engineering schools, able to combine design thinking with synthetic knowledge bases typical of technology activities. A second kind concerns 'aesthetic design': those activities are common in soft innovation processes where designers focus on the creation of new products and new meanings, hence combine design knowledge with symbolic knowledge. These designers are more likely to be trained in arts schools or dedicated design schools. The last row also includes the metrics that we will use for each activity, which we will explain in detail in the data section. Next, we move to explain how regional specializations can stem from all three activities.

Regional innovation specializations and the principle of relatedness

As already hinted in the previous section, innovation activities tend to develop in a path-dependent manner and the opportunities for further diversification or specialization get shaped by the knowledge and capabilities already developed at each point in time (Boschma, 2017). This intuition from evolutionary economics has been conceptualized into the 'principle of relatedness' (Hidalgo *et al.*, 2007) and extensively applied within evolutionary economic geography. When it comes to innovation activities, researchers have provided strong evidence for the significance of relatedness in shaping the emergence of new regional technological specializations (Kogler et al. 2013; Boschma et al. 2015; Petralia et al. 2017; Apa et al. 2018). The rationale is that regions will branch to new technologies that are related to their existing technological capabilities by tapping into and recombining existing knowledge base. The underlying mechanism behind related diversification relies on the idea that related pieces of knowledge and capabilities are easier to be recombined thanks to cognitive proximity (Rigby, 2015). At the same time knowledge spillovers is not the only mechanism supporting relatedness. As discussed by Boschma (2017) regions can show specialization in the same two activities because of knowledge spillovers, skill relatedness or input-output relations: hence, relatedness can stem both from similarity and complementarity.

Scholars have used the principle of relatedness to examine branching in several economic activities in addition to technological ones, including export products (e.g. Hidalgo et al. 2007), industries (e.g. Neffke et al. 2011) and scientific fields (e.g. Boschma et al. 2014). Key to our arguments in this paper is that the underlying logic of new knowledge emerging from the recombination of related bits of existing knowledge appears indeed to apply to different kinds of knowledge, not only technological one. To illustrate how the logic is helpful to conceptualize developments in all the three knowledge spaces of technology, design and market, we refer to an example. Smart phone technologies recombine technologies related to batteries, chips, antennas, audio, video, display and the Internet (Castaldi et al. 2015). At the same time, one can view the corresponding product, i.e. the smart phone, as defining a new product category that recombines communication devices, photographic instruments, fashion items and recreational services (Suarez et al. 2015). Similarly, smart phone product developers have experimented with different design options, often working around trade-offs in the size, power and portability of the new devices (Cecere et al. 2015).

What this example illustrates is how knowledge spillovers exist for all knowledge types. Within technology spaces, technological similarity or complementarity will tend to support cognitive proximity. Within design spaces, design options where regions tend to co-specialize can be seen as closer to each other because of similar or complementary types of design solutions. With product spaces, similar or complementary symbolic knowledge will make regional specializations in two given product categories more likely than in two categories without a common or connected meaning.

If the principle of relatedness can indeed apply to all three innovation activities, then we can derive hypotheses about the emergence of new regional specializations in design and market activities that are similar to those about technology specializations. Specifically, we expect regions to be more likely to develop new innovation specializations when they show a high degree of relatedness of knowledge in each specific space.

From relatedness to cross-relatedness between innovation activities

The discussion above has treated the three innovation activities as independent ones. Yet, there are many ways in which innovation activities are related, through formal input-output relations but also knowledge feedback loops and even skill relatedness. The three knowledge spaces are also likely to be strongly connected with each other, as opting for specific technologies often comes with restricted design and market choices, and the other way around. These linkages allow expanding the notion of relatedness to include 'cross-relatedness' as well.

We are not the first to extend the principle of relatedness to more knowledge dimensions. Catalan et al. (2020) focused on how scientific capabilities of a nation can contribute to new related technology specializations and defined a 'sci-tech space'. Pugliese et al. (2019) leveraged a complexity perspective to investigate multi-layered networks of relations between science, technology and product capabilities of countries. Our work differs in several respects. First, our interest is in the regional level. Research has demonstrated that both national and regional systems of innovation are important, but the regional level allows capturing variance in innovation activities that is left unexplained when taking a national lens (Cooke et al. 1997). Second, our focus goes beyond the scientific base of regions and concerns instead on the more applied stages of innovation, those mostly happening within corporate borders. Pugliese et al. (2019) did include downstream activities by considering new product specializations, using export data. Such data are not focused on innovation and also tend to underestimate service activities, which are harder to trade. Finally, both studies assume a linear relation from science to technology and then market. Instead, our framework can accommodate cross-relatedness to run both ways, from upstream to downstream but also the other way around.

Based on our characterization of the three innovation activities, we propose to conceptualize relatedness between activities, i.e. cross-relatedness, as the co-occurrence of specialization in two different innovation activities as revealed by the patterns at the supra-regional level (whole Europe or whole US, in our case).

Cross-relatedness and regional innovation specializations

We discuss here the mechanisms behind cross-relatedness that we expect to play a role. Similarly to the understanding of relatedness, the mechanisms may be quite diverse: while it is hard to discriminate them empirically (Boschma, 2017), we can discuss the ones that are most likely to be at play. In the first place, co-occurrence of two focal innovation activities, say design and market ones, in one region can be there because local companies possess knowledge that is useful for both activities, hinting at synergies in the underlying learning processes (Farinha et al. 2019). In the second place, there might be more formal input-output relations which connect innovation activities from technology to market (Essletzbichler, 2015).

Starting with the ideas of the linear model of technology-push innovation, one could expect clear patterns of cross-relatedness of technology to the downstream innovation activities. From a different perspective, Chan et al. (2017) show how technological advances can push the boundaries of designs and even alter the styles of entire product segments. In general, one would expect regions co-specializing in technology and technological design activities to rely on similar or complementary knowledge bases (Corradini and Karoglou, 2022). Such *tech-design relatedness* would likely stem from underlying synergies in technology-driven innovation (Murmann and Frenken, 2006; Dan et al. 2018) and mostly concern technical design. In such processes, *tech-market relatedness* would also be there, as working combinations of synthetic and symbolic knowledge bases feed co-specialization in specific technologies and specific markets (Hise et al., 1989; Breznitz, 2021).

On the other hand, demand-pull arguments might also be mechanisms for cross-relatedness between innovation activities. Firms may experience synergies from embedding technological advances into their branded products/services as a way to deal with increased competition (Greenhalgh and Rogers, 2012). For instance, Bei (2019) showed how firms may source technology from other firms to capitalize on their already successful brands. There is also evidence that clusters of firms with strong market positions have incentives to invest in new technological specializations to keep their products up to speed with technological upgrades (Fritsch and Wyrwich, 2021). This suggests synergies that would support *market-tech relatedness* too.

As for cross-relatedness between design and market activities, mechanisms supporting regional co-specialization can also run both ways. They are likely to be strongest when aesthetic design is concerned, relying on synergies between design and symbolic knowledge. In fact, design knowledge specialization can be leading in creating synergies with specific market knowledge bases (Walsh 1996; d'Ippolito, 2014), but there is also rich evidence of significant feedback loops that design activities rest upon, with market demand or user feedback shaping new product aesthetics and functionalities (Di Stefano et al. 2012). These mechanisms are likely to be more evident in the case of aesthetic design and for industries where soft innovation works as a key competitive advantage.

Let us also stress two main reasons *not* to expect cross-relatedness. A first reason might be that the three innovation activities are independent or to a large extent separable. Within global or even simply modularized value chains, inventions can occur in one place and commercialization in another. If this is the case, then our framework would pick-up the resulting lack of cross-relatedness and demonstrate it for those specific type of activities for which indeed separation is possible. A second reason is that downstream innovation activities may not need technological inventions to capitalize from. In several low-tech sectors, innovation rests upon 'soft' elements or gets prompted by user feedback: this is the case in many service sectors but also in the creative and cultural industries (Stoneman, 2010; Schmoch and Gauch, 2009; Millot, 2009).

Relatedness, cross-relatedness and regional innovation specializations

Our framework suggests that regional innovation specializations in each innovation stage can be explained by both relatedness in that activity and cross-relatedness with the other two activities. Table 2 illustrates this in a matrix form, where the diagonal elements are the relatedness elements expected to be positively associated with each regional innovation specialization (the three column headers). The goal is to elaborate on which relatedness and cross-relatedness dimensions we expect to play a more pronounced role for the different innovation specializations. We do so by considering the two types of design, which will also correspond to the two empirical contexts where we test our hypotheses.

For all three innovation specializations we expect the relevant relatedness measure (i.e. the diagonal elements in Table 2) to reveal the strongest association with the emergence of new specialization, following prior theoretical and empirical literature.

For technological specialization we also expect cross-relatedness with downstream activities to play a role, but less than relatedness. We expect tech-design relatedness to be positively associated with new technological specialization mostly when design is of a technical nature. Here the underlying synergies between technological and design knowledge appear more evident.

For design specializations, we expect differences for technical design vs aesthetic design, with a stronger role for tech-design cross-relatedness, leveraging clear synergies between synthetic and design knowledge bases vs a stronger role for design-market relatedness, leveraging clear synergies between design and symbolic knowledge bases.

For market specializations, we similarly expect market relatedness to reveal the strongest association, but cross-relatedness with more upstream activities should also significantly matter. We envision a role for design-market relatedness, in case of aesthetic design, and for tech-market relatedness, overall.

IPRs as innovation proxies

To capture technological inventions, design and market activities, we consider utility patents, industrial designs and trademarks. The main advantage of opting for these three metrics is that they are all intellectual property rights, hence comparable types of data, with some common strengths and limitations. Key common strengths are that IPRs can be counted at regional and national level and they are registered after undergoing a formal filing procedure where specific requirements are checked. Key common limitations are at least two. First, their validity as innovation metrics is weakened by strategic practices in their filings (Greenhalgh and Rogers, 2010). One way to take into account this problem is to consider only filings that made it to registration, which allows disregarding at least some of the strategic practices. This is the approach we opt for in our analysis. Second, IPRs only measure a share of all activities that contribute to innovation. However, they do capture activities that add value to the economy. At the regional level, there are several studies that relate at least one of these IPRs and sometimes more than one, to innovation and/or entrepreneurship (Torres-Preciado et al., 2014; Mendonça., 2014; Drivas 2020; Corradini and Karoglou, 2022; Pinate et al. 2022) while Filippetti et al. (2019) showed that regions engaging in all three types of IPR activity appeared more economically resilient.

Utility patents have been employed most extensively as innovation metric, to capture technological inventions and the upstream phase of innovation processes (Griliches 1990). While researchers have critiqued the patent system, as several inventions may not pass the patentability threshold (Bessen and Meurer, 2008; Boldrin and Levine, 2013), scholars have also shown that patents can provide financial incentives to inventors, hence are specifically used by actors more strongly investing in invention (Moser 2005; Lerner 2009).

Design rights protect the aesthetics of industrial products and have been discussed as potential metrics for innovation (Stoneman, 2010, Filitz et al. 2015, Filippetti and d'Ippolito, 2017; Heikkilä and Peltoniemi, 2019). Through the use of design rights, researchers have shown the evolution of products and styles in an entire industry (Chan et al. 2017). All prior studies have focused either on the US or Europe, but a key difference between the two systems is that in the US they are actually design patents, while in Europe design rights are more similar to trademarks (Schickl, 2013). As such, in the US design rights undergo a similar procedure to patents and are tested for novelty and industrial applicability, while design rights in Europe capture new designs that fulfill the condition of distinctiveness. This institutional difference allows connecting US design patents primarily to new technical designs and technology-based innovation processes and European design rights primarily to new aesthetic designs typical of industries focused on soft innovation. Hence, it also allows testing for hypotheses involving technical design in the US context and those involving aesthetic design in the European context.

Trademarks are distinctive signs that protect differentiating attributes of a product or service (Graham *et al.*, 2013). Empirical studies have found significant evidence that trademarks correlate positively to innovation activity and new product/service introduction (Mendonça et al. 2004; Flikkema et al. 2014, 2019). What distinguishes trademarks from the other IPRs is that the applicant needs to provide evidence of use in commerce before being granted (Graham et al. 2013, Schautschick and Greenhalgh, 2016). Hence, trademarks can capture the most downstream stage of innovation activity.

Utility patents can be classified according to the International Patent Classification (IPC). Design rights are classified by the international Locarno classification of design categories: these categories concern industrial design and connect to specific artefacts. Therefore, these Locarno categories have an intuitive connection with both patent classes, since patents have to indicate an industrial application, and trademark classes, since they indicate specific product categories that identify the specific markets where trademark owners claim protection. Trademark classes are defined by the international Nice classification, including 45 classes (1-34 cover goods and 35-45 cover services). A strength of combining these classifications is that the three are internationally comparable. A limitation is that they differ in the degree of detail, with patent classes being the most detailed, followed by design and trademark classes.

Methods

Data collection

Starting with Europe, we collected utility patents filed at the EPO from the OECD's REGPAT database (Maraut et al. 2008), including applicant's NUTS-2 information, and trademark and industrial designs filed at the EUIPO. Each EUIPO record was separately located in an xml file and after a careful reconfiguration we obtained each applicant's country and postal code information. We assigned each postal code to a NUTS-2 region based on European Commission's NUTS-2 postal codes concordance.[1] We included a country's NUTS-2 regions if more than 90% of the country's trademarks and industrial designs were assigned to a NUTS-2 region. The countries that did not satisfy this criterion were: Bulgaria, Ireland, Romania and Lithuania.[2] Further, we dropped countries that only included a single NUTS-2 region as in Xiao et al. (2018): Estonia, Cyprus, Luxembourg, Latvia and Malta. Finally, we included Switzerland and Norway though they are not part of the EU, due to the proximity to other EU countries. Overall, we obtained information for approximately 450 thousand patents,

[1] https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units/nuts.
[2] For Bulgaria, Ireland and Romania we could not locate the postal code for approximately 50% of the trademark filings from the database. For Lithuania, while we could locate a postal code for the trademark application data, we could only obtain a three-digit postal code. However, from the European Commission's NUTS-2 postal codes concordance we could only locate a five-digit postal code thereby excluding this country due to the lack of clear concordance.

640 thousand industrial designs and 570 thousand trademarks, filed during 2003-2016, that made it to registration and spread over 259 NUTS-2 regions.

For the US case, we collected utility patent, design patent and trademark records from the public databases of USPTO.[3] In line with other studies, we chose the MSA level as the geographical level most comparable to NUTS-2 regions (Lee and Rodriguez-Pose, 2013).[4] For each USPTO record, we obtained the application and registration dates, the associated classes and the location information of the applicant. Note that patents can also be counted by inventor location but since trademarks can only be counted by owner location we opt for the applicant location for all IPRs. For both utility and design patents, the USPTO has already geocoded the applicants, hence we only needed to assign the coordinates to MSAs based on US Census' TIGER shapefiles, 2010 version. For the case of trademarks, we used the postal codes and assigned them to MSAs based on the same shapefiles. Missing postal codes were searched in Google Earth Pro and based on their coordinates were once again assigned to an MSA. Overall, we obtained information for approximately 1 million utility patents, 137 thousand design patents, and 3.6 million trademarks, filed during the period 2003-2016 that made it to registration, and assigned to 363 MSAs.

For utility patents, the standard practice in the evolutionary economic geography literature is to focus on either the 3-digit (Balland et al. 2019) or 4-digit IPC classification (Montresor and Quatraro 2017). Given the higher level of detail we opt for the first listed 4-digit IPC classification. For design rights, we employ the 4-digit classification Locarno classification. For trademark classes, we rely on the 45 Nice classes, but we discuss possible extensions in the robustness tests and conclusions.

Finally, we decided to count IPRs by filing year, i.e. the year closest to the underlying activity taking place, but we only counted registered IPRs. This allows counting IPR filings that underwent the administrative checks of the formal requirements and hence are likely to be of higher quality than filings that did not make it to registration. This choice also explains why our sample ends in 2016. Because of the known lags between filing and registration, we can exploit more recent information to check registration. Note that EUIPO started accepting industrial design applications only in 2003; therefore, to provide an even comparison across both testbeds, our samples start in 2003.

Key variables

Our approach relies on calculating relatedness measures for all three innovation metrics. In doing so, we first build technology, design and market spaces and then a comprehensive innovation space, where the three activities are related to each other. For patents and design rights we consider the main primary listed class. For the case of trademarks, we opt for whole counting in case that there is more than one Nice class disclosed. Unlike patents, and design rights, for an applicant to claim an additional class s/he needs to provide evidence that the trademark is used in commerce in all selected classes of goods/services. Therefore, a trademark with several classes has a wider scope of commercial activity compared to a trademark with a single class. Nonetheless, we run several robustness checks where we opt for alternative choices of the listed classes.

We discuss here the construction of the innovation spaces for the European case and provide examples specific to this context. The analysis for the US follows exactly the same

[3] Office of the Chief Economist: https://www.uspto.gov/ip-policy/economic-research/research-datasets. For a thorough overview of this trademark database, see Graham et al. (2013).

[4] Lee and Rodríguez-Pose (2013) focused on MSAs and NUTS-1 regions when comparing US and European regions. However, wherever data availability allowed them to use NUT-2 regions, they performed the analysis at that level.

steps. First, we bundle years in three time periods: 2003-2008, 2009-2012 and 2013-2016. Therefore, the period dimension, denoted as *t* takes three values: *t=0* for 2003-2008, *t=1* for 2009-2012 and for and *t=2* for 2013-2016. Working with periods instead of single years ensures that a region's entry into a new specialization is robust and not due to a random short-term shock (Neffke et al. 2011). We first constructed the indicator of specialization that has become standard in the relatedness literature, inspired by Balassa (1965). The indicator identifies whether region *r* has a Revealed Comparative Advantage (RCA) in class *i* for a particular IPR during period *t*. For instance, for market specializations, the RCA is defined as:

$$RCA_{r,i,t} = \frac{trademarks_{r,i,t}/\sum_i trademarks_{r,i,t}}{\sum_r trademarks_{r,i,t}/\sum_r \sum_i trademarks_{r,i,t}}$$

In other words, $RCA_{r,i,t}$ at period *t* measures the share of trademarks in class *i* that region *r* filed over the share of trademarks filed in class *i* of all trademarks filed. Therefore, a higher $RCA_{r,i,t}$ implies that region *r* is relatively more active in trademark class *i* compared to the entire set of regions. Similar specialization indicators are calculated for patents and design rights. To have a class index which runs through all three classifications, we re-code class *i* to be a numeric index that takes values between 1-836 (1-822) for Europe (US). This comes from the fact that for Europe (US), we have 589 (591) 4-digit IPC classes and 202 (186) Locarno classes and 45 Nice classes.

Following the literature (Hidalgo et al. 2007), a region is specialized in class *i* if its RCA is above one:

$$x_{r,i,t} = \begin{cases} 1 \; if \; RCA_{r,i,t} > 1 \\ 0 \quad otherwise \end{cases}$$

The next step generates the key inputs for constructing the innovation spaces capturing the underlying relatedness and cross-relatedness. Following the literature, we start from estimating proximities among classes from revealed patterns of co-specialization. We calculate the probability that a region specializes in class *i* given that it also specializes in class *j*. For the 259 NUTS-2 regions (or 363 MSAs) we count the instances where class *i* has an RCA>1 given that class *j*, where *i≠j*, has RCA>1. Then by dividing this number with the instances where class *j* has an RCA>1 we obtain the probability $P(x_{i,t}|x_{j,t})$. This probability does not need to be equal to the opposite conditional probability$P(x_{j,t}|x_{i,t})$. To reconcile this asymmetric distance between classes, we follow Hausmann and Klinger (2007) and calculate the minimum of each pair of probabilities. That is:

$$\varphi_{i,j} = \min\{P(x_{i,t}|x_{j,t}), P(x_{j,t}|x_{i,t})\}$$

For the European (US) case, $\varphi_{i,j}$ populate an *836x836*(*822x822*) matrix of proximities which capture the overall innovation space. Note that this matrix is symmetric by definition, as $\varphi_{i,j} = \varphi_{j,i}$ for each given combination of *i* and *j*.

Figures A1 and A2 of the Online Appendix display the innovation space for Europe and US respectively. After summing all φ's we calculate the Minimum Spanning Tree (MST) algorithm to display the edges between nodes. For both geographical contexts, we observe several clusters where technology, design and market activities are interconnected. For the case of Europe trademark market classes are linked with many technology classes within a core cluster comprised of loosely connected smaller clusters. In the case of the US the picture is slightly different with trademark market classes dispersed across the innovation space instead of within a central cluster. Let us provide some examples that illustrate how the specific patterns of relatedness and cross-relatedness have mattered for shaping new regional specializations. In Europe, specializations in industrial designs on locking and closing devices tend to co-occur with designs in chain links and permanent magnets, but also with patents on bolts, hinges and devices for opening and closing any type of wing. The DE11 region (Stuttgart) displayed a new specialization in locking and closing devices after it had developed specializations in both the related design and patent fields. In US, the MSA of Tampa-St. Petersburg-Clearwater, FL exhibited a new market specialization in clothes and footwear trademarks after it already specialized in patents that include inventions in outerwear, protective garments and accessories.

Going beyond these specific examples, our analysis aims at establishing to what extent regional relatedness and cross-relatedness matter on average for the emergence of new regional technology, design and market specializations. To this end, we estimate regression models that allow us to gauge the strength and directionality of relatedness within and between the three innovation activities on the emergence of new regional specializations. The dependent variables capture the entry of region *r* in a new specialization in a particular class *i*. They take the value of 1 if region *r* exhibits an RCA in period *t* given it had not in period *t-1* and 0 otherwise. That is:

$$Entry_{r,i,t} = \begin{cases} 1\ if\ x_{r,i,t} = 1\ \ and\ x_{r,i,t-1} = 0 \\ \quad 0 \quad otherwise \end{cases}$$

We then construct the main independent variables of interest, capturing regional relatedness within and between types of IPRs. Following the literature, we use average density measures, as they are called in the literature, which consider proximities of the focal class to the classes where the region is already specialized in. For exposition, assume that our interest is on $Entry_{r,i,t}$ where i=1-45, that is, we focus on new market specializations. We construct three variables. The first one captures relatedness specific to market activities:

$$Market\_RELATEDNESS_{i,r} = \frac{\sum_{j=1, j\in r, j\neq i}^{45} \varphi_{ij}}{\sum_{j=1, j\neq i}^{45} \varphi_{ij}}$$

The numerator is the sum of φ_{ij} in the trademark classes *j* that region *r* is specialized in. The denominator is the overall sum of φ_{ij} for market class *i*. This measure captures how embedded trademark market class *i* is in the rest of the regional market activities.

The other two variables, capturing cross-relatedness, are:

$$TechMarket\_RELATEDNESS_{i,r} = \frac{\sum_{j=46, j\in r, j\neq i}^{634} \varphi_{ij}}{\sum_{j=46, j\neq i}^{634} \varphi_{ij}}$$

$$DesignMarket\_RELATEDNESS_{i,r} = \frac{\sum_{j=635, j\in r, j\neq i}^{836} \varphi_{ij}}{\sum_{j=635, j\neq i}^{836} \varphi_{ij}}$$

These two variables capture the relatedness of class *i* to technology and design classes where the region is also specialized in.

Similar variables are constructed for the models explaining the emergence of new technological and design specializations. Note that cross-related density measures are not symmetric: $TechMarket$ cross-related density is different from $MarketTech$ related density, given that the variables depend on the regional specializations.

Econometric specifications

To examine the role of the different relatedness measures for new specializations, we consider separate regressions for each innovation specialization. Each regression can be understood as the operationalization of the relations in the three columns of Table 2, which summarizes our conceptual framework.

For new market specializations we consider:

$$Entry_{r,i,t} = \alpha_0 + \alpha_1 Market\_RELATEDNESS_{r,i,t-1} + \alpha_2 TechMarket\_RELATEDNESS_{r,i,t-1} + \alpha_3 DesignMarket\_RELATEDNESS_{r,i,t-1} + RegionPeriod_{r,t} + ClassPeriod_{i,t} + \varepsilon_{r,i,t} \quad (1)$$

For new technology specializations:

$$Entry_{r,i,t} = \beta_0 + \beta_1 Tech\_RELATEDNESS_{r,i,t-1} + \beta_2 MarketTech\_RELATEDNESS_{r,i,t-1} + \beta_3 DesignTech\_RELATEDNESS_{r,i,t-1} + RegionPeriod_{r,t} + ClassPeriod_{i,t} + \varepsilon_{r,i,t} \quad (2)$$

For new design specializations:

$$Entry_{r,i,t} = \gamma_0 + \gamma_1 Design\_RELATEDNESS_{r,i,t-1} + \gamma_2 MarketDesign\_RELATEDNESS_{r,i,t-1} + \gamma_3 TechDesign\_RELATEDNESS_{r,i,t-1} + RegionPeriod_{r,t} + ClassPeriod_{i,t} + \varepsilon_{r,i,t} \quad (3)$$

Note that the first period's (2003-2008) information is utilized as lagged information for the period 2009-2012. To this end, we can only observe entries in periods 2009-2012 and 2013-2016. Overall, we expect relatedness to be positively related to new specializations and hence $\alpha_1 > 0$, $\beta_1 > 0$ and $\gamma_1 > 0$. In addition, we expect cross relatedness measures to be positively associated to new specializations too (i.e. $\alpha_2 > 0$, $\beta_2 > 0$, $\gamma_2 > 0$, $\alpha_3 > 0$, $\beta_3 > 0$ and $\gamma_3 > 0$). To be able to compare coefficients across regressions as well as interpreting them, all relatedness measures are standardized as in Xiao et al. (2018). Note that we only include region-class-period observations where the region did not display an RCA above 1 in period *t-1* (i.e. $x_{r,i,t-1} = 0$). If the region had already specialized in that class, then that region-class pair would add no information on the relation between the relatedness measures and new specializations. To take into account region and class intertemporal heterogeneity, we include both *region-period* and *class-period* fixed effects in all regressions. Due to the large amount of fixed effects all regressions are estimated via OLS since non-linear estimators, such as probit and logit, are likely to produce biased estimates (Greene 2012; Boschma et al. 2013; Gomila 2020). Standard errors are clustered at the region-class level to avoid serial correlation (Bertrand et al. 2004).

Empirical analysis

Descriptive and graphical analysis

Table A1 of the Online Appendix shows summary statistics of the dependent variables. For the European (US) case, there is a 13% (15%) probability that a new specialization will take place in a region-trademark class, while for patents and designs the probabilities are 6% (5%) and 7% (5%) respectively. The lower likelihood of technology and design specializations is to be expected given the larger number of technology and design classes as compared to the market classes. Also, we observe that the likelihood of new specializations is similar in the two geographical contexts. Tables A2-A4 show the correlations of the dependent and independent variables for Europe (Panel A) and US (Panel B). Relatedness measures correlate strongly, which might result in a multicollinearity bias in the econometric analysis. To examine whether multicollinearity confounds the overall empirical results, we always include relatedness and cross-relatedness variables stepwise in the regressions.

Figures 1A-1C display the average relatedness measures for the period 2003-2008 for European regions. Both average $Market\_RELATEDNESS$ and $Design\_RELATEDNESS$ are not always high (low) in regions with high (low) average relatedness in technological activities. The difference between market and design relatedness on one hand and technology relatedness on the other hand suggests that the first follow own dynamics which might be (at least partly) independent of technological ones. This also corroborates our intuition that EUIPO's designs relate to aesthetic design activities closer to market activities than technological ones. Figure 2A-2C display the average $Market\_RELATEDNESS$, $Tech\_RELATEDNESS$ and $Design\_RELATEDNESS$ for US MSAs. Similarly to the European case, the map of market relatedness reveals a somewhat different pattern than technological relatedness. However, unlike the European case, regions tend to score similarly in terms of technology and design relatedness and high scores often coincide with regions with strong technological profiles. This seems in line with the fact that USPTO design rights capture technical design activities.

Regression results

In what follows we present the baseline results for the market, design and technology specializations, i.e. equations (1)-(3) referring to the three columns of Table 2. In the three tables, we include the results for Europe and US. Starting with the more downstream innovation activity, we analyse the emergence of new market specializations in products and services (Table 3). Note that the Variance Inflation Factor (VIF) for the case of Europe when all coefficients are included (Column 3) $Market\_RELATEDNESS$, $TechMarket\_RELATEDNESS$ and $DesignMarket\_RELATEDNESS$ are 1.18, 2.15 and 2.09 while for the US (Column 6) the VIF tests are 1.12, 3.6 and 3.47. While these measures could be considered high, they are far from the critical threshold of 10 (Hair et al. 2009). Also, the stepwise inclusion of the relatedness measures (Columns 1-3 and 4-6) does not reveal any dramatic change in the coefficients, hence multicollinearity does not appear to be an issue.

For both geographical contexts, relatedness is significantly associated to new market specializations. As the independent variables are standardized, the coefficient in the full model for Europe (Column 3) can be interpreted as follows: a one standard deviation increase of $Market\_RELATEDNESS_{r,i,t-1}$ from its mean is associated to an increase in the likelihood that region r will exhibit a new specialization in market class i of 22.0 percentage units. $TechMarket\_RELATEDNESS_{r,i,t-1}$ is also strongly associated to new market specializations. A result that stands out is that

$DesignMarket\_RELATEDNESS$ is neither positive nor significant for the US case. We go back to this finding after presenting all baseline results.

Table 4 reports estimates for the role of relatedness in the emergence of new technological specializations. The VIF for $Tech\_RELATEDNESS$, $MarketTech\_RELATEDNESS$ and $DesignTech\_RELATEDNESS$ are 2.39, 1.24 and 2.32 for EU (Column 3), while for US (Column 6) they are 4.18, 1.14 and 4.11 respectively. Also in this case, multicollinearity is not an issue. $Tech\_RELATEDNESS$ is strongly associated to new technological specializations, in both contexts. Cross-relatedness measures are significant, but the coefficients are significantly lower than the $Tech\_RELATEDNESS$ coefficient (t-tests comparisons are statistically significant at $p<0.01$), indicating that relatedness matters more than cross-relatedness when it comes to new technology specializations.

Table 5 displays estimates for the role of relatedness in new design specializations. The VIF for $Design\_RELATEDNESS$, $MarketDesign\_RELATEDNESS$ and $TechDesign\_RELATEDNESS$ is 2.29, 1.22 and 2.35 for EU (Column 3) while for US (Column 6) 4.03, 1.13 and 4.10 respectively. Once again, the stepwise inclusion of the variables does not reveal any multicollinearity issues. All relatedness measures are significantly and positively related to the emergence of new design specializations. For the US, cross-relatedness matters even more than relatedness, particularly when it comes to technology (t-test comparisons statistically significant at $p<0.01$).

We can now relate our findings back to the framework and hypotheses we proposed in Table 2. We refer to the baseline results in Tables 3, 4 and 5, as the empirical counterparts of the three columns in Table 2, with the US context offering a testbed for the case of technical design and Europe for aesthetic design. We found that relatedness mattered the most in all regressions, in line with our expectations. There were two exceptions. For US regions tech-design cross-relatedness was as strongly associated with new design specializations as design relatedness, pointing to technical design specializations being strongly driven by technology. Also, for European regions, new market specializations were equally strongly associated to all relatedness and cross-relatedness measures, suggesting strong synergies between all three innovation activities supporting market leadership.

A consistent pattern across both geographical contexts was that cross-relatedness with technology mattered when considering new market and design specializations, to an extent comparable to relatedness and in line with the relevance of synergies from upstream to downstream activities. When focusing on new technology specializations, cross-relatedness with market and design mattered but much less so than technology relatedness, as expected for backward linkages from downstream to upstream.

Going back to the results for market specializations (Table 3), we noted that $DesignMarket\_RELATEDNESS$ played no role in the case of US while it exhibited a strong positive coefficient in the case of Europe. To further compare this finding to our intuition that indeed *DesignMarket* relatedness would mostly be there in technology-driven innovation processes, we also checked whether results changed when focusing on high-tech market specializations only. We focused on a subset of trademark classes which can be related to high technology products, as suggested by Mendonça and Fontana (2011)[5] and then estimate the same regressions. Column 2 of Table 6 shows that for high tech product market specializations, the coefficient of $DesignMarket\_RELATEDNESS$ for the US is indeed positive and significant. For consistency with the baseline, we also run a similar estimation for the European case (Column 1). Overall, the results confirm the positive coefficients for all relatedness and cross-relatedness measures, suggesting strong synergies between all innovation activities for market specializations in high-tech products.

[5] These high technology product Nice classes are 1, 3, 5, 7 and 9-15.

Robustness checks

To provide robustness checks for the above results we consider several variations. Table A5 in the online appendix provides an overview of all the robustness results, to highlight similarities and spot instances where the results deviate from the baseline regressions. First, while we opted for OLS to include an array of fixed effects and control for unobserved heterogeneity, we wish to check that the choice of this estimator is not driving our results. To this end, we estimated all regressions via probit models too. We had to drop class-period dummies due to convergence issues, but Table A6 reveals similar results to the baseline estimates. Note that the small change in sample size from the baseline results in the probit estimations comes from the fact that a few observations are predicted perfectly and hence are excluded.

Further, for the European case we had excluded five countries (EE, CY, LT, LV and MT) since they included a single NUTS-2 region. In Table A7 we add these five countries as additional regions. Results are again similar to those of the baseline models.

We also validated several critical choices we made about our IPR metrics. A first issue is whether assigning patents to the inventor location, instead of the applicant location, makes a difference. Inventor location is available for USPTO and EPO patents and for USPTO design patents, while all trademarks and EUIPO design rights only include the applicant location. This is a well-known issue when combining patent and trademark data. There might be a headquarters effect, with regions hosting headquarters scoring higher on those activities measured with trademarks simply because trademark filings only include information on the applicant company. Yet it should be noted that activities related to marketing and commercialization tend to be more centralized at headquarter level than upstream research activities anyways (Castaldi and Mendonça, 2022). In Table A8, we report results of the models after repeating all analyses with inventor locations for those IPRs where they are available. Results are comparable to the baseline ones.

Further, we estimated all regressions with variables calculated using all IPR filings, not only those that made it to registration. By doing so, we are including more filings, whose quality might be lower. Focusing on filings might be interesting for two reasons: first, it provides a more timely indicator, given that registration takes some time; second, it includes activities by companies that did not have the required financial resources or the expertise to obtain a successful registration. After re-estimating the relatedness variables, Table A9 provides the counterparts of Tables 3, 5 and 6. Overall, results are quite similar.

We also validated our choices in terms of counting classes. On one hand, we considered fractional counting of trademarks instead of whole counting. After reconstructing all the variables, we perform the same analysis. Results are displayed in Columns 1 and 4 of Tables A10-A12 for each new innovation specialization for the US and Europe.

On the other hand, a known limitation of trademark classes is that they are only 45, thereby potentially underestimating new trademark specializations, and also affecting any relatedness measure associated with the trademarks. While we cannot provide a direct robustness check with alternative versions of trademark classes, we can provide an indirect test. We considered technology classes (123 in total) at the 3-digit IPC classification instead of the 4-digit level, making the level of detail of patent classes more similar to trademark classes. In this case the $\varphi_{i,j}$ matrix for Europe (US) populates a *370x370* (*354x354*) matrix. If the aggregation for the 45 Nice classes was problematic, then this analysis would deliver starkly different results since the level of aggregation for IPC classes also changed dramatically. In Columns 2 and 5 of Tables A10-A12 we re-estimated the baseline models by whole counting trademarks while in Columns 3 and 6 we opted for fractional counting. In both cases we considered the 3-digit IPC classification. The results are by and large similar to the baseline ones. Finally, we revisited the choice of considering only the first-listed IPC class.[6] We considered all the 4-digit IPC classes for each utility patent and Table A13 reports the alternative results. They are qualitatively similar with the exception of $DesignTech\_RELATEDNESS$, whose coefficient for new technological specialization of US regions not significant. Yet, the main result that relatedness has a stronger association than cross-relatedness remains.

Conclusion

In this paper we aimed to take seriously the calls from researchers and policymakers for a broader view on regional innovation specializations, beyond technology only. We developed a conceptual framework grounded on combining insights from innovation studies and evolutionary economic geography. In our framework we conceptualized three main types of innovation activities and argued that the principle of relatedness can be leveraged to understand branching to new specializations within and between the three innovation activities.

We also showed how IPR metrics can be used to capture developments in three underlying knowledge spaces of technology/patents, design/design rights and markets/trademarks. Our empirical analysis of US and EU regions in three recent periods provided support for an overall strong association of both relatedness and cross-relatedness measures with the emergence of new regional innovation specializations. This confirmed that path-dependence and place-dependence act as powerful force in technological, design and market trajectories. At the same time, we found that cross-relatedness played a significant role in the emergence of new regional specializations for all three innovation activities. Design appeared as an intermediary function lying in between the two other innovation activities and intertwined with both, albeit in different ways. The two geographical testbeds helped us to gauge the role of technical and aesthetic design activities. We found design-market cross-relatedness to matter for new market specializations in the European context, while that link was only there for high-tech product market specialization in the US case.

To expand further on policy implications, our results can inform the development and implementation of regional policies of smart specialization in several ways. First, considering more downstream specializations appears relevant, since actual innovation that has reached the commercialization phase is important to generate jobs and entrepreneurial opportunities in regions. In fact, the latest take on S3 smart specialization strategies (European Commission, 2021) acknowledges that: "Social, organisational, market and service innovation, or practice-based innovation, play as important a role in S3 as technological innovation based on scientific research" (p.2). As Foray et al. (2011) put it, with reference to cases like the one of Pierre-Hyacinthe Caseaux: "the outcome of the process is much more than a "simple" technological innovation..." resulting in a new activity that offers to the regional economy "…superior commercial prospects." (p. 6). Additionally, our analysis can be seen as complementary to approaches focused on the roots and upstream drivers of innovation specializations, specifically concerned with the development of regional scientific strongholds (Catalan et al. 2020).

Second, our analysis has demonstrated that regions have different strengths in each innovation stage and a focus on technology only overshadows opportunities for regions that do not belong to the small circle of high-tech clusters. Some regions may exploit a history of related design and market capabilities to uncover further specializations even without investing

[6] We did not pursue a similar robustness for design rights. Design patents in the USPTO data simply did not include any secondary Locarno class, while the EUIPO data only included it for 3.5% of the filings. Therefore, focusing on the first-listed Locarno class is the only option for the overwhelming majority of the data.

in technology (Breznitz, 2021). In practice, policymakers can analyse innovation spaces to uncover patterns of co-specialization along the innovation process. They can draw much more fine-grained maps than what we could show, by leveraging the public and timely innovation metrics that we suggested here. Even though regional and national innovation scoreboards (like the EU Regional Innovation Scoreboard and the Science and Engineering indicators of the US National Science Foundation) by now include trademark and design rights counts next to patent ones, such aggregate counts can hardly characterize regional specializations in a qualitative manner and help to uncover specific strengths and weaknesses. Instead, unleashing the richness of the information on technology, design and market classes where local companies are filing different IPRs allows mapping opportunities and challenges of smart specialization strategies through a relatedness lens.

Our study offers the potential for several extensions and validation exercises. A key limitation of our analysis was the coarseness of the trademark classes, which allowed us to capture market relatedness only between very broad product categories. Ongoing efforts to define more granular subclasses using text analysis of goods and service descriptions will offer the opportunity to work at the same level of detail of patent classes (Neuhäusler et al. 2021; Abbasiharofteh et al. 2022). This will allow to better align empirics with the conceptual interpretation of relatedness in the market space. Another research direction would be to validate our results using alternative metrics. For instance, trademarks could be substituted with trade data, in line with how Hidalgo et al. (2007) and Pugliese et al. (2019) capture the product space. Trademark activity is likely to be related to export activity, especially when considering registrations at supra-national offices like the EUIPO. Yet, trademark data also capture specializations in non-tradable activities (mostly low-tech services) which will not be covered by trade data. These activities might not matter directly for innovation, still a comparison of patterns could be interesting.

Finally, our focus on NUTS-2 regions is not without limitations. There is a perennial issue noted as modifiable areal unit problem, pointing at the fact that performing the same analysis on smaller geographic units could reveal non-trivial differences (Fotheringham and Wong, 1991).Yet, an additional problem that would arise is the presence of too many zeros in the IPR metrics. A similar argument could be made for focusing on MSAs for the case of US instead of counties or cities. Comparative analysis of different geographical levels could reveal significant insights on the implied spillovers of Smart Specialization policies to larger areas, an issue that has only recently been examined in the literature (Balland and Boschma, 2021).

Finally, we envision the potential for several extensions of our framework. One extension could be to move beyond overall average patterns and analyse heterogeneity in how relatedness and cross-relatedness matter, for instance for economically developed ones vs lagging regions. This would align with work suggesting that the explanatory power of relatedness differs by region type (e.g. Petralia et al. 2017). Alternatively, different IPR filings in urban vs rural regions might also be at play and could be controlled for (Wojan, 2019). Finally, one could use our innovation space approach to zoom in on specific innovation specializations that might be particularly desirable from a strategic or societal perspective. For instance, future research could extend the rich literature on green technology specializations (e.g. Barbieri et al. 2020) and look at relatedness dynamics for regional green innovation specializations beyond technology (in line with the firm-level analysis in Ghisetti et al., 2021). Similarly, one could focus on the digital revolution and analyse the extent to which regions might specialize in technology, design or market activities related to artificial intelligence or Industry 4.0. Ultimately, this goes in the direction of pushing for a broader take on regional innovation capabilities and the policies that can support them.

References

Abbasiharofteh, Milad, Carolina Castaldi, and Sergio Petralia. 2022. "From patents to trademarks: towards a concordance map." Final report EPO Academic Research Program, https://www.epo.org/learning/materials/academic-research-programme/research-project-grants.html

Apa, Roberta, Ivan De Noni, Luigi Orsi, and Silvia Rita Sedita. 2018. "Knowledge space oddity: How to increase the intensity and relevance of the technological progress of European regions." *Research Policy* 47 (9):1700-1712.

Asheim, Bjørn T, and Lars Coenen. 2005. "Knowledge bases and regional innovation systems: Comparing Nordic clusters." *Research Policy* 34 (8):1173-1190.

Balassa, Bela. 1965. "Trade liberalisation and "revealed" comparative advantage." *The manchester school* 33 (2):99-123.

Balland, Pierre-Alexandre, and Ron Boschma. 2021. "Complementary interregional linkages and Smart Specialisation: An empirical study on European regions." *Regional Studies* 55 (6):1059-1070.

Barbieri, Nicoló, François Perruchas, and Davide Consoli. 2020. "Specialization, diversification, and environmental technology life cycle." *Economic Geography* 96 (2):161-186.

Bei, Xiaoshu. 2019. "Trademarks, specialized complementary assets, and the external sourcing of innovation." *Research Policy*.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119 (1):249-275.

Bessen, J., and M. J. Meurer. 2008. *Patent failure: How Judges, Bureaucrats, and Lawyers put Innovators at Risk*: Princeton University Press.

Boldrin, Michele, and David K. Levine. 2013. "The Case Against Patents." *Journal of Economic Perspectives* 27 (1):3-22.

Boschma, Ron. 2017. "Relatedness as driver of regional diversification: A research agenda." *Regional Studies* 51 (3):351-364.

Boschma, Ron, Gaston Heimeriks, and Pierre-Alexandre Balland. 2014. "Scientific knowledge dynamics and relatedness in biotech cities." *Research policy* 43 (1):107-114.

Boschma, Ron, Pierre-Alexandre Balland, and Dieter Franz Kogler. 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.

Breznitz, Dan. 2021. *Innovation in real places: Strategies for prosperity in an unforgiving world*: Oxford University Press, USA.

Capello, Roberta, and Camilla Lenzi. 2013. "Territorial patterns of innovation: a taxonomy of innovative regions in Europe." *The Annals of regional science* 51 (1):119-154.

Castaldi, Carolina, Koen Frenken, and Bart Los. 2015. "Related variety, unrelated variety and technological breakthroughs: an analysis of US state-level patenting." *Regional studies* 49 (5):767-781.

Castaldi, Carolina, and Sandro Mendonça. 2022. "Regions and trademarks: research opportunities and policy insights from leveraging trademarks in regional innovation studies." *Regional Studies* 56 (2):177-189.

Catalán, Pablo, Carlos Navarrete, and Felipe Figueroa. 2020. "The scientific and technological cross-space: Is technological diversification driven by scientific endogenous capacity?" *Research Policy*:104016.

Cecere, Grazia, Nicoletta Corrocher, and Riccardo David Battaglia. 2015. "Innovation and competition in the smartphone industry: Is there a dominant design?" *Telecommunications Policy* 39 (3-4):162-175.

Chan, Tian Heong, Jürgen Mihm, and Manuel E Sosa. 2017. "On Styles in Product Design: An Analysis of US Design Patents." *Management Science* 64 (3):1230-1249.

Cooke, Philip, Mikel Gomez Uranga, and Goio Etxebarria. 1997. "Regional innovation systems: Institutional and organisational dimensions." *Research policy* 26 (4-5):475-491.

Corradini, Carlo, and Michail Karoglou. 2022. "On the persistence and complementarities of design and technological change: a regional perspective." *Regional Studies*:1-14.

Dan, Sujan M, Brian I Spaid, and Charles H Noble. 2018. "Exploring the sources of design innovations: Insights from the computer, communications and audio equipment industries." *Research Policy* 47 (8):1495-1504.

Davids, Mila, and Koen Frenken. 2018. "Proximity, knowledge base and the innovation process: Towards an integrated framework." *Regional studies* 52 (1):23-34.

Di Stefano, Giada, Alfonso Gambardella, and Gianmario Verona. 2012. "Technology push and demand pull perspectives in innovation studies: Current findings and future research directions." *Research policy* 41 (8):1283-1295.

Dosi, Giovanni. 2000. "Opportunities, incentives and the collective patterns of technological change." In *Innovation, Organization and Economic Dynamics*, 145-162. Edward Elgar Publishing.

Drivas, Kyriakos. 2022. "The role of technology and relatedness in regional trademark activity." *Regional Studies* 56 (2):242-255.

d'Ippolito, Beatrice. 2014. "The importance of design for firms' competitiveness: a review of the literature." *Technovation*:forhcoming.

Essletzbichler, Jürgen. 2015. "Relatedness, industrial branching and technological cohesion in US metropolitan areas." *Regional Studies* 49 (5):752-766.

European Commission. 2021. Key elements of smart specialization strategies, Accessed at https://s3platform.jrc.ec.europa.eu/documents/portlet_file_entry/20125/S3-Key-Elements.pdf/23a14b4c-f871-9a77-7e93-0b19e4b910f1

Farinha Fernandes, Teresa, Pierre-Alexandre Balland, Andrea Morrison, and Ron Boschma. 2019. "What drives the geography of jobs in the us? unpacking relatedness." *Industry and Innovation*:1-35.

Feldman, Maryann P. 1994. *The geography of innovation*. Vol. 2: Springer Science & Business Media.

Filippetti, Andrea, and Beatrice D'Ippolito. 2017. "Appropriability of design innovation across organisational boundaries: exploring collaborative relationships between manufacturing firms and designers in Italy." *Industry and Innovation* 24 (6):613-632.

Filippetti, Andrea, Petros Gkotsis, Antonio Vezzani, and Antonio Zinilli. 2019. "How to survive an economic crisis? Lessons from the innovation profiles of EU regions." *Working Paper. European Commission.*

Filitz, Rainer, Joachim Henkel, and Bruce S Tether. 2015. "Protecting aesthetic innovations? An exploration of the use of registered community designs." *Research Policy* 44 (6):1192-1206.

Flikkema, Meindert, Carolina Castaldi, Ard-Pieter de Man, and Marcel Seip. 2019. "Trademarks' relatedness to product and service innovation: A branding strategy approach." *Research Policy*.

Flikkema, MJ, AP de Man, and C Castaldi. 2014. "Are Trademark Counts a Valid Indicator of Innovation? Results of an In-Depth Study of New Benelux Trademarks Filed by SMEs." *Industry and Innovation* 21 (4):310-331.

Foray, D., David, P. A., & Hall, B. H. 2009. Smart specialisation: The concept*. In E. Commission (Ed.), Knowledge for growth: Prospects for science, technology and Innovation* (pp. 25–29). Brussels: European Commission. DG Research.

Foray, D., David, P. A., & Hall, B. H. 2011. "Smart specialisation from academic idea to political instrument, the surprising career of a concept and the difficulties involved in its implementation." *(No. REP_WORK). EPFL.*

Foray, Dominique, Kevin Morgan, and Slavo Radosevic. 2018. "The role of smart specialisation in the EU research and innovation policy landscape." *European Commission*.

Fritsch, Michael, and Michael Wyrwich. 2021. "Does successful innovation require large urban areas? Germany as a counterexample." *Economic Geography* 97 (3):284-308.

Gemser, Gerda, and Mark AAM Leenders. 2001. "How integrating industrial design in the product development process impacts on company performance." *Journal of Product Innovation Management* 18 (1):28-38.

Ghisetti, Claudia, Sandro Montresor, and Antonio Vezzani. 2021. "Design and environmental technologies: Does 'green-matching'actually help?" *Research Policy* 50 (5):104208.

Gomila, Robin. 2021. "Logistic or linear? Estimating causal effects of experimental treatments on binary outcomes using regression analysis." *Journal of Experimental Psychology: General* 150 (4):700.

Graham, Stuart J. H., Galen Hancock, Alan C. Marco, and Amanda Fila Myers. 2013. "The USPTO Trademark Case Files Dataset: Descriptions, Lessons, and Insights." *Journal of Economics & Management Strategy* 22 (4):669-705.

Greene, William. 1994. "Accounting for excess zeros and sample selection in Poisson and negative binomial regression models." *NYU Working Paper No. EC-94-10 (1994)*

Greenhalgh, Christine, and Mark Rogers. 2012. "Trade marks and performance in services and manufacturing firms: evidence of Schumpeterian competition through innovation." *Australian Economic Review* 45 (1):50-76.

Greenhalgh, Christine, and Mark Rogers. 2010. "Innovation, intellectual property, and economic growth." In *Innovation, Intellectual Property, and Economic Growth*. Princeton University Press.

Griliches, Zvi. 1990. "Patent statistics as economic indicators: a survey." *National Bureau of Economic Research*

Hair Jr, Joseph F, Wiliam C Black, Barry J Babin, and Rolph E Anderson. 2010. "Multivariate data analysis." In *Multivariate data analysis*, 785-785.

Hausmann, Ricardo, and Bailey Klinger. 2007. The structure of the product space and the evolution of comparative advantage. *CID Working Paper Series.*

Heikkilä, Jussi, and Mirva Peltoniemi. 2019. "Great expectations: Learning the boundaries of design rights." *Research Policy*:103795.

Hidalgo, César A, Bailey Klinger, A-L Barabási, and Ricardo Hausmann. 2007. "The product space conditions the development of nations." *Science* 317 (5837):482-487.

Hidalgo, César A, Pierre-Alexandre Balland, Ron Boschma, Mercedes Delgado, Maryann Feldman, Koen Frenken, Edward Glaeser, Canfei He, Dieter F Kogler, and Andrea Morrison. 2018. "The principle of relatedness." International conference on complex systems.

Hise, Richard T, Larry O'Neal, James U McNeal, and Ananthanarayanan Parasuraman. 1989. "The effect of product design activities on commercial success levels of new industrial products." *Journal of product innovation management* 6 (1):43-50.

Kogler, Dieter F, David L Rigby, and Isaac Tucker. 2013. "Mapping knowledge space and technological relatedness in US cities." *European Planning Studies* 21 (9):1374-1391.

Lee, Neil, and Andrés Rodríguez-Pose. 2013. "Innovation and spatial inequality in Europe and USA." *Journal of economic geography* 13 (1):1-22.

Lerner, Josh. 2009. "The Empirical Impact of Intellectual Property Rights on Innovation: Puzzles and Clues." *American Economic Review Papers and Proceedings* 99 (2):343-348.

Leten, Bart, Rene Belderbos, and Bart Van Looy. 2007. "Technological diversification, coherence, and performance of firms." *Journal of Product Innovation Management* 24 (6):567-579.

Marques, Pedro, and Kevin Morgan. 2018. "The heroic assumptions of smart specialisation: A sympathetic critique of regional innovation policy." *New avenues for regional innovation systems-theoretical advances, empirical cases and policy lessons*:275-293.

Marsili, Orietta, and Ammon Salter. 2006. "The dark matter of innovation: design and innovative performance in Dutch manufacturing." *Technology analysis & strategic management* 18 (5):515-534.

Maraut, Stéphane, Hélène Dernis, Colin Webb, Vincenzo Spiezia, and Dominique Guellec. 2008. The OECD REGPAT Database. OECD Publishing.

Mendonça, Sandro, Tiago Santos Pereira, and Manuel Mira Godinho. 2004. "Trademarks as an indicator of innovation and industrial change." *Research Policy* 33 (9):1385-1404.

Mendonça, Sandro, and Roberto Fontana. 2011. Estudo sobre o Contributo das Marcas para o Crescimento Económico e para a Competitividade Internacional. INPI.

Mendonça, Sandro. 2014. "National adaptive advantages: Soft innovation and marketing capabilities in periods of crisis and change." In *Structural Change, Competitiveness and Industrial Policy*, 149-166. Routledge.

Millot, Valentine. 2009. "Trademarks as an indicator of product and marketing innovations." *Paris: OECD Publishing* (No. 2009/6).

Moser, Petra. 2005. "How Do Patent Laws Influence Innovation? Evidence from Nineteenth-Century World's Fairs." *American Economic Review* 95 (4):1214-1236.

Murmann, Johann Peter, and Koen Frenken. 2006. "Toward a systematic framework for research on dominant designs, technological innovations, and industrial change." *Research policy* 35 (7):925-952.

Neffke, Frank, Martin Henning, and Ron Boschma. 2011. "How do regions diversify over time? Industry relatedness and the development of new growth paths in regions." *Economic Geography* 87 (3):237-265.

Neuhäusler, Peter, Alexander Feidenheimer, Rainer Frietsch, and Henning Kroll. 2021. *Generating a classification for EUIPO trademark filings: A string matching approach*: Fraunhofer ISI.

Petralia, Sergio, Pierre-Alexandre Balland, and Andrea Morrison. 2017. "Climbing the ladder of technological development." *Research Policy* 46 (5):956-969.

Pinate, Adriana C, Alessandra Faggian, Claudio Di Berardino, and Carolina Castaldi. 2022. "The heterogenous relationship between migration and innovation: Evidence from Italy." *Industry and Innovation*:1-25.

Pugliese, Emanuele, Giulio Cimini, Aurelio Patelli, Andrea Zaccaria, Luciano Pietronero, and Andrea Gabrielli. 2019. "Unfolding the innovation system for the development of countries: coevolution of Science, Technology and Production." *Scientific reports* 9 (1):1-12.

Radosevic, S. 2018. "Fostering innovation in less-developed and low institutional capacity regions: Challenges and opportunities." Background paper for an OECD/EC Workshop on.

Rigby, David L. 2015. "Technological relatedness and knowledge space: entry and exit of US cities from patent classes." *Regional Studies* 49 (11):1922-1937.

Rodríguez-Pose, Andrés, and Neil Lee. 2020. "Hipsters vs. geeks? Creative workers, STEM and innovation in US cities." *Cities* 100:102653.

Santamaría, Lluís, María Jesús Nieto, and Andrés Barge-Gil. 2009. "Beyond formal R&D: Taking advantage of other sources of innovation in low-and medium-technology industries." *Research Policy* 38 (3):507-517.

Schautschick, Philipp, and Christine Greenhalgh. 2016. "Empirical studies of trade marks–the existing economic literature." *Economics of innovation and new technology* 25 (4):358-390.

Schickl, Lena. 2013. "Protection of industrial design in the United States and in the EU: Different concepts or different labels?" *The Journal of world intellectual property* 16 (1-2):15-38.

Schmoch, Ulrich, and Stephan Gauch. 2009. "Service marks as indicators for innovation in knowledge-based services." *Research Evaluation* 18 (4):323-335.

Schumpeter, Joseph A. 1934. *The theory of economic development: an inquiry into profits, capital, credit, interest, and the business cycle*. Edited by Harvard University Press. Cambridge, MA.

Stoneman, Paul. 2010. *Soft innovation: economics, product aesthetics, and the creative industries*: Oxford University Press.

Suarez, Fernando F, Stine Grodal, and Aleksios Gotsopoulos. 2015. "Perfect timing? Dominant category, dominant design, and the window of opportunity for firm entry." *Strategic Management Journal* 36 (3):437-448.

Torres-Preciado, Víctor Hugo, Mayrén Polanco-Gaytán, and Miguel Á Tinoco-Zermeño. 2014. "Technological innovation and regional economic growth in Mexico: a spatial perspective." *The Annals of Regional Science* 52:183-200.

Verganti, Roberto. 2006. "Innovating through design." *Harvard business review* 84 (12):114.

Verganti, Roberto. 2008. "Design, meanings, and radical innovation: A metamodel and a research agenda." *Journal of product innovation management* 25 (5):436-456.

Walsh, Vivien. 1996. "Design, innovation and the boundaries of the firm." *Research Policy* 25 (4):509-529.

Windrum, Paul, Koen Frenken, and Lawrence Green. 2017. "The importance of ergonomic design in product innovation. Lessons from the development of the portable computer." *Industrial and Corporate Change* 26 (6):953-971.

Wojan, Timothy R. 2019. "Geographical differences in intellectual property strategies and outcomes: Establishment-level analysis across the American settlement hierarchy." *Regional Studies, Regional Science* 6 (1):574-595.

Xiao, Jing, Ron Boschma, and Martin Andersson. 2018. "Industrial diversification in Europe: The differentiated role of relatedness." *Economic Geography*:1-36.

Table 1: Three key innovation activities and their properties.

Innovation activity	**Technology**	**Design**	**Market**
Main output	Technological inventions	Novel designs	New products (goods and services)
Phase of the innovation process	Research	Design and Prototyping	Product Development and Marketing
Type of knowledge	Technological, synthetic, engineering-based	a. Technical design, in technology-driven innovation b. Aesthetic design, in soft innovation	Symbolic knowledge, categories, meanings
Knowledge space	Technological space	Design space	Market space
Proxy/Metric	Patents	a. Design patents b. Designs	Trademarks

Table 2: Relatedness and cross-relatedness behind regional innovation specializations: expected strength of relationships depending on type of design activity

	Regional innovation specializations		
Dimensions of (cross)relatedness	**Technology specialization**	**Technical Design specialization**	**Market specialization**
Technology	Technological relatedness (+++)	Tech-design relatedness (++)	Tech-market relatedness (++)
Technical Design	Design-tech relatedness (+)	Design Relatedness (+++)	Design-market relatedness (+)
Market	Market-tech relatedness (+)	Market-design relatedness (+)	Market Relatedness (+++)

Dimensions of (cross)relatedness	**Technology specialization**	**Aesthetic Design specialization**	**Market specialization**
Technology	Technological relatedness (+++)	Tech-design relatedness (+)	Tech-market relatedness (++)
Aesthetic Design	Design-tech relatedness (+)	Design Relatedness (+++)	Design-market relatedness (++)
Market	Market-tech relatedness (+)	Market-design relatedness (+)	Market Relatedness (+++)

Table 3: Role of relatedness and cross-relatedness for new market specializations.

	Europe			US		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
$Market\_RELATEDNESS_{r,i,t-1}$	0.228***	0.229***	0.220***	0.242***	0.248***	0.242***
	(0.012)	(0.012)	(0.012)	(0.015)	(0.015)	(0.015)
$TechMarket\_RELATEDNESS_{r,i,t-1}$	0.158***		0.134***	0.137***		0.141***
	(0.041)		(0.041)	(0.043)		(0.044)
$DesignMarket\_RELATEDNESS_{r,i,t-1}$		0.171***	0.150***		0.002	-0.021
		(0.041)	(0.041)		(0.042)	(0.043)
Constant	-0.126*	-0.176***	0.014	-0.036	0.136**	-0.026
	(0.071)	(0.068)	(0.071)	(0.065)	(0.060)	(0.068)
Observations	14,336	14,336	14,336	21,011	21,011	21,011
R-squared	0.129	0.129	0.130	0.090	0.089	0.090
adj R-Squared	0.090	0.091	0.091	0.053	0.053	0.053

Note: The dependent variable in all regressions is $Entry_{r,i,t}$. All regressions are estimated via ordinary least squares (OLS) and include region–period and Nice class-period dummies. Columns (1)-(3) consider the European case while Columns (4)-(6) consider the US case. Standard errors are clustered at the region–class level and are displayed in parentheses. ***p<0.01; **p<0.05; *p<0.1.

Table 4: Role of relatedness and cross-relatedness for new technology specializations.

	Europe			US		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
$Tech\_RELATEDNESS_{r,i,t-1}$	0.194***	0.191***	0.188***	0.187***	0.181***	0.181***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
$MarketTech\_RELATEDNESS_{r,i,t-1}$	0.011***		0.011***	0.005***		0.005***
	(0.001)		(0.001)	(0.001)		(0.001)
$DesignTech\_RELATEDNESS_{r,i,t-1}$		0.020***	0.019***		0.011***	0.012***
		(0.003)	(0.003)		(0.003)	(0.003)
Constant	-0.068***	0.046***	0.046***	-0.051***	-0.061***	-0.058***
	(0.010)	(0.008)	(0.008)	(0.012)	(0.012)	(0.012)
Observations	268,815	268,815	268,815	379,247	379,247	379,247
R-squared	0.089	0.089	0.089	0.115	0.115	0.115
adj R-Squared	0.083	0.083	0.083	0.110	0.110	0.110

Note: The dependent variable in all regressions is $Entry_{r,i,t}$. All regressions are estimated via ordinary least squares (OLS). All columns include region–period and IPC class-period dummies. Columns (1)-(3) consider the European case while Columns (4)-(6) consider the US case. Standard errors are clustered at the region–class level and are displayed in parentheses. ***p<0.01; **p<0.05; *p<0.1.

Table 5: Relatedness and cross-relatedness for new design specializations.

	Europe			US		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
$Design\_RELATEDNESS_{r,i,t-1}$	0.094***	0.086***	0.084***	0.088***	0.056***	0.056***
	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
$MarketDesign\_RELATEDNESS_{r,i,t-1}$	0.017***		0.015***	0.008***		0.006***
	(0.002)		(0.002)	(0.002)		(0.002)
$TechDesign\_RELATEDNESS_{r,i,t-1}$		0.053***	0.049***		0.099***	0.098***
		(0.006)	(0.006)		(0.005)	(0.005)
Constant	0.085***	0.077***	0.005	0.052**	0.062***	0.065***
	(0.021)	(0.021)	(0.021)	(0.021)	(0.018)	(0.018)
Observations	91,686	91,686	91,686	123,493	123,493	123,493
R-squared	0.081	0.081	0.082	0.110	0.114	0.114
adj R-Squared	0.072	0.072	0.073	0.102	0.106	0.106

Note: The dependent variable in all regressions is $Entry_{r,i,t}$. All regressions are estimated via ordinary least squares (OLS). All columns include region–period and Locarno class-period dummies. Columns (1)-(3) consider the European case while Columns (4)-(6) consider the US case. Standard errors are clustered at the region–class level and are displayed in parentheses. ***p<0.01; **p<0.05; *p<0.1.

Table 6: The case of high-tech product market specializations.

	Europe	US
VARIABLES	(1)	(2)
$Market\_RELATEDNESS_{r,i,t-1}$	0.225***	0.185***
	(0.030)	(0.034)
$TechMarket\_RELATEDNESS_{r,i,t-1}$	0.606***	0.289***
	(0.145)	(0.108)
$DesignMarket\_RELATEDNESS_{r,i,t-1}$	0.188*	0.174*
	(0.109)	(0.097)
Constant	-0.042	-0.118
	(0.167)	(0.099)
Observations	3,608	5,145
R-squared	0.221	0.181
adj R-Squared	0.084	0.041

Note: The dependent variable in all regressions is $Entry_{r,i,t}$. Both columns consider Nice classes that are related to medium and high technology industries according to Mendonça and Fontana (2011). These Nice classes are 1, 3, 5, 7 and 9-15. All regressions are estimated via ordinary least squares (OLS). All columns include region–period dummies and Nice class-period dummies. Standard errors are clustered at the region–class level and are displayed in parentheses.

Figure 1A Technology relatedness for the period 2003-2008.

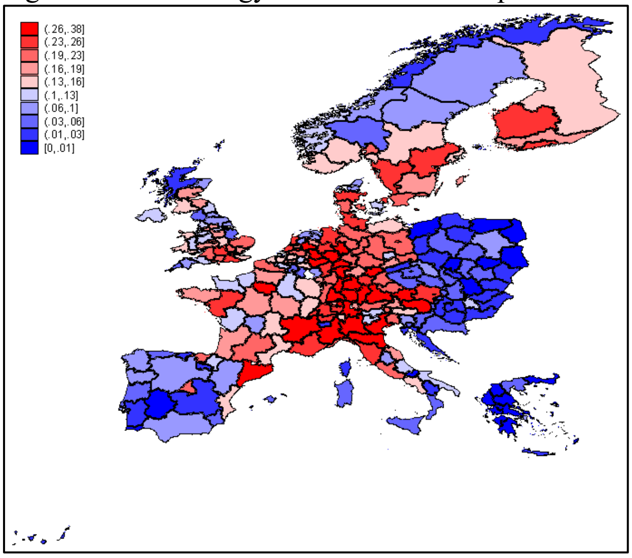


Figure 1B. Design relatedness for the period 2003-2008.

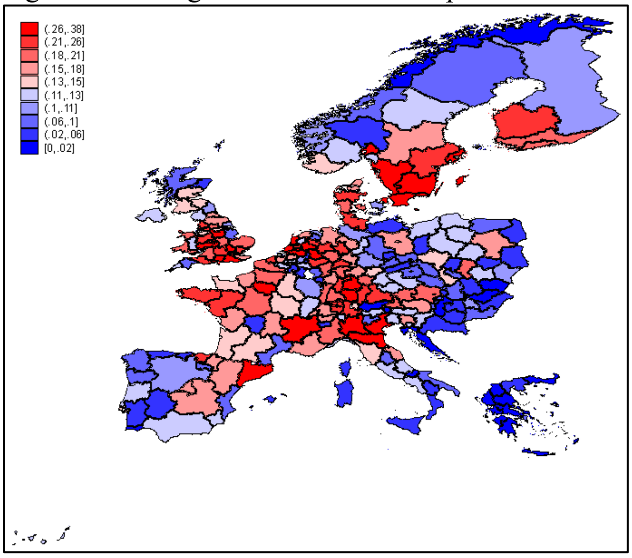


Figure 1C. Market relatedness for the period 2003-2008.

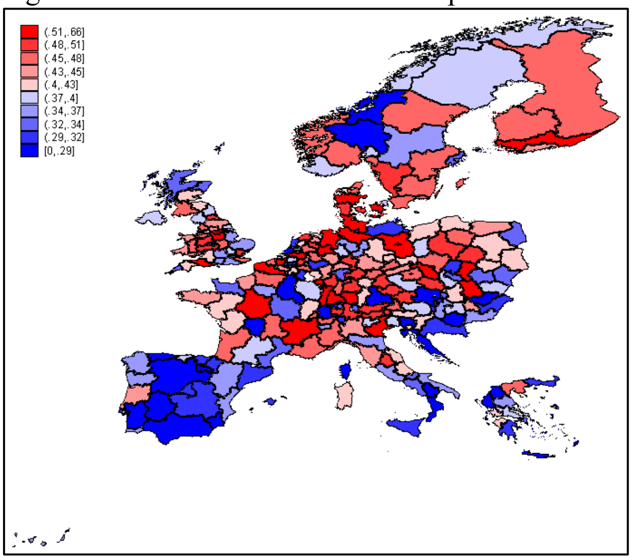


Figure 2A. Technology relatedness for the period 2003-2008.

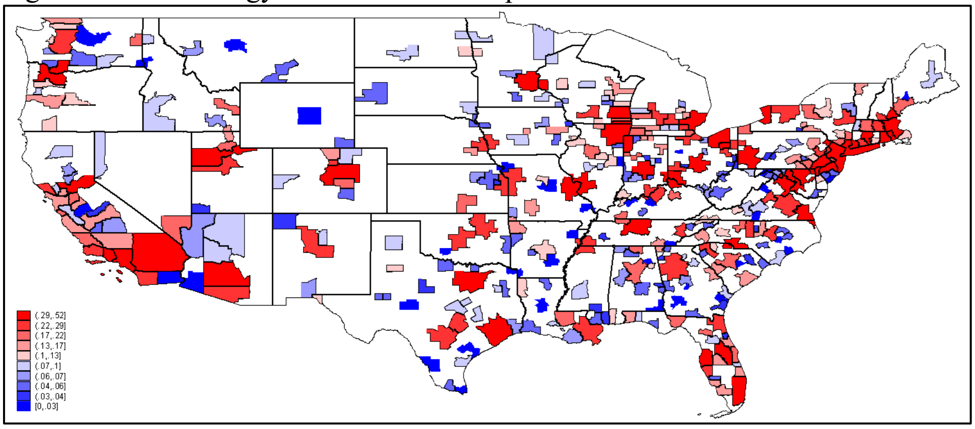


Figure 2B. Design relatedness for the period 2003-2008.

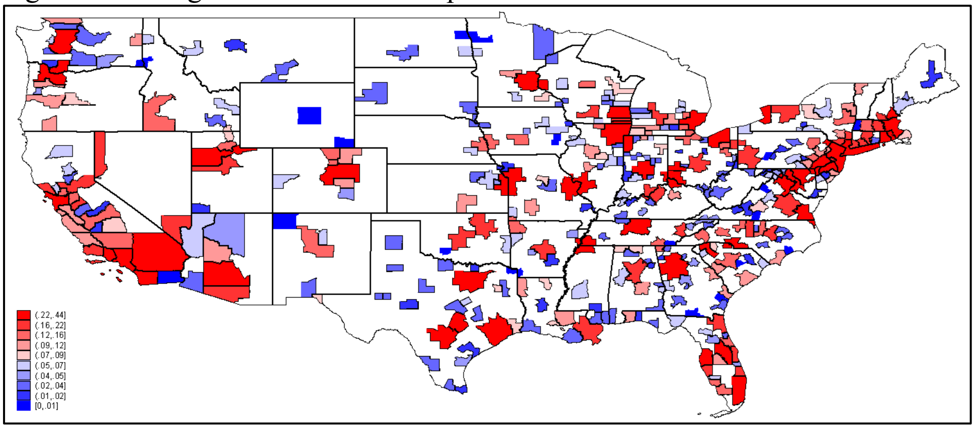
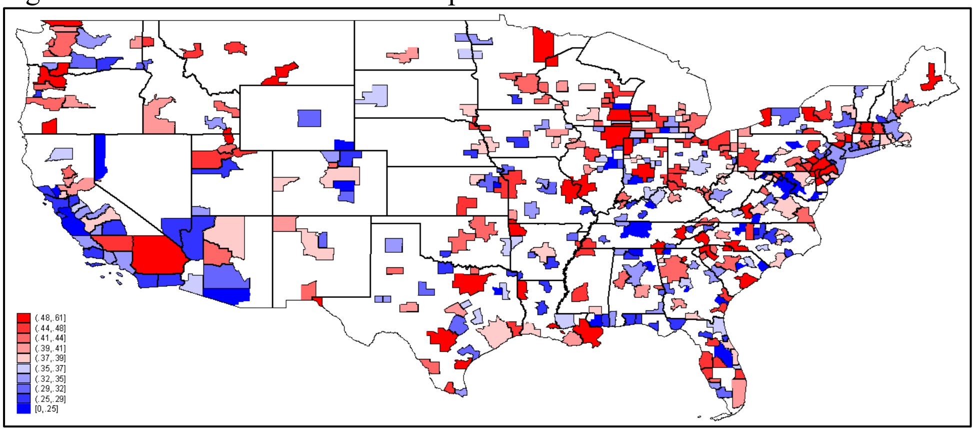


Figure 2C. Market relatedness for the period 2003-2008.



Online Appendix

Table A1: Summary statistics: means (standard deviations)

	European Data		
	Trademark Classes (n=23,310)	Patent classes (n=305,102)	Design classes (n=104,636)
$Entry_{r,i,t}$	0.13	0.06	0.07
	(0.34)	(0.23)	(0.26)
$x_{r,i,t-1}$	0.38	0.12	0.12
	(0.49)	(0.32)	(0.33)
	US Data		
	Trademark Classes (n=32,670)	Patent classes (n=426,888)	Design classes (n=135,036)
$Entry_{r,i,t}$	0.15	0.05	0.05
	(0.35)	(0.22)	(0.22)
$x_{r,i,t-1}$	0.36	0.11	0.09
	(0.48)	(0.31)	(0.28)

Table A2: Correlation matrices for trademark activity.

Panel A. European Data.

	$Entry_{r,i,t}$	$x_{r,i,t-1}$	$Market\_RELATEDNESS_{r,i,t-1}$	$TechMarket\_RELATEDNESS_{r,i,t-1}$	$DesignMarket\_RELATEDNESS_{r,i,t-1}$
$Entry_{r,i,t}$	1				
$x_{r,i,t-1}$	-0.31	1			
$Market\_RELATEDNESS_{r,i,t-1}$	0.02	0.28	1		
$TechMarket\_RELATEDNESS_{r,i,t-1}$	-0.05	0.15	0.40	1	
$DesignMarket\_RELATEDNESS_{r,i,t-1}$	-0.04	0.14	0.36	0.71	1

Panel B. US Data.

	$Entry_{r,i,t}$	$x_{r,i,t-1}$	$Market\_RELATEDNESS_{r,i,t-1}$	$TechMarket\_RELATEDNESS_{r,i,t-1}$	$DesignMarket\_RELATEDNESS_{r,i,t-1}$
$Entry_{r,i,t}$	1				
$x_{r,i,t-1}$	-0.31	1			
$Market\_RELATEDNESS_{r,i,t-1}$	0.01	0.14	1		
$TechMarket\_RELATEDNESS_{r,i,t-1}$	-0.05	0.11	0.35	1	
$DesignMarket\_RELATEDNESS_{r,i,t-1}$	-0.06	0.09	0.30	0.85	1

Note: All correlations are significant at the 1% level.

Table A3: Correlation matrices for patent activity.

Panel A. European Data.

	$Entry_{r,i,t}$	$x_{r,i,t-1}$	$Tech\_RELATEDNESS_{r,i,t-1}$	$MarketTech\_RELATEDNESS_{r,i,t-1}$	$DesignTech\_RELATEDNESS_{r,i,t-1}$
$Entry_{r,i,t}$	1				
$x_{r,i,t-1}$	-0.09	1			
$Tech\_RELATEDNESS_{r,i,t-1}$	0.16	0.40	1		
$MarketTech\_RELATEDNESS_{r,i,t-1}$	0.07	0.20	0.45	1	
$DesignTech\_RELATEDNESS_{r,i,t-1}$	0.12	0.34	0.77	0.41	1

Panel A. US Data.

	$Entry_{r,i,t}$	$x_{r,i,t-1}$	$Tech\_RELATEDNESS_{r,i,t-1}$	$MarketTech\_RELATEDNESS_{r,i,t-1}$	$DesignTech\_RELATEDNESS_{r,i,t-1}$
$Entry_{r,i,t}$	1				
$x_{r,i,t-1}$	-0.08	1			
$Tech\_RELATEDNESS_{r,i,t-1}$	0.18	0.42	1		
$MarketTech\_RELATEDNESS_{r,i,t-1}$	0.07	0.18	0.38	1	
$DesignTech\_RELATEDNESS_{r,i,t-1}$	0.16	0.37	0.88	0.35	1

Note: All correlations are significant at the 1% level.

Table A4: Correlation matrices for design activity.

Panel A. European Data.

	$Entry_{r,i,t}$	$x_{r,i,t-1}$	$Design\_RELATEDNESS_{r,i,t-1}$	$MarketDesign\_RELATEDNESS_{r,i,t-1}$	$TechDesign\_RELATEDNESS_{r,i,t-1}$
$Entry_{r,i,t}$	1				
$x_{r,i,t-1}$	-0.10	1			
$Design\_RELATEDNESS_{r,i,t-1}$	0.12	0.34	1		
$MarketDesign\_RELATEDNESS_{r,i,t-1}$	0.06	0.16	0.39	1	
$TechDesign\_RELATEDNESS_{r,i,t-1}$	0.10	0.29	0.76	0.42	1

Panel A. US Data.

	$Entry_{r,i,t}$	$x_{r,i,t-1}$	$Design\_RELATEDNESS_{r,i,t-1}$	$MarketDesign\_RELATEDNESS_{r,i,t-1}$	$TechDesign\_RELATEDNESS_{r,i,t-1}$
$Entry_{r,i,t}$	1				
$x_{r,i,t-1}$	-0.07	1			
$Design\_RELATEDNESS_{r,i,t-1}$	0.20	0.38	1		
$MarketDesign\_RELATEDNESS_{r,i,t-1}$	0.08	0.17	0.35	1	
$TechDesign\_RELATEDNESS_{r,i,t-1}$	0.20	0.37	0.88	0.37	1

Note: All correlations are significant at the 1% level.

Table A5. Overview of regressions results, both baseline and robustness ones: + signs based on significant statistical differences (t-tests comparison with $p<0.05$) between relatedness and cross-relatedness coefficients in each model.

Europe (aesthetic design case)

	Theory	B	High-Tech Products	P	Add5	I	F	C1	C2	C3	M
					New technology specializations						
$Tech\_RELATEDNESS$	(+++)	(+++)		(+++)	(+++)	(+++)	(+++)	(+++)	(+++)	(+++)	(+++)
$MarketTech\_RELATEDNESS$	(+)	(+)		(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)
$DesignTech\_RELATEDNESS$	(+)	(+)		(++)	(+)	(+)	(+)	(+)	(+)	(+)	(+)
					New design specializations						
$Design\_RELATEDNESS$	(+++)	(+++)		(+++)	(+++)	(+++)	(+++)	(+++)	(+++)	(+++)	(+++)
$TechDesign\_RELATEDNESS$	(+)	(+)		(++)	(+)	(+++)	(++)	(++)	(++)	(++)	(++)
$MarketDesign\_RELATEDNESS$	(+)	(+)		(+)	(+)	(+)	(+)	(+)	(+)	(+)	(+)
					New market specializations						
$Market\_RELATEDNESS$	(+++)	(+++)	(++)	(+++)	(+++)	(+++)	(+++)	(+++)	(+++)	(+++)	(+++)
$DesignMarket\_RELATEDNESS$	(++)	(+++)	(++)	(+++)	(+++)	(+++)	(+++)	(+++)	(+++)	(+++)	(+++)
$TechMarket\_RELATEDNESS$	(++)	(+++)	(+++)	(++)	(+++)	(+++)	(++)	(++)	(++)	(++)	(+++)

US (technical design case)

	Theory	B	High-Tech Products	P	Add5	I	F	C1	C2	C3	M
					New technology specializations						
$Tech\_RELATEDNESS$	(+++)	(+++)		(+++)		(+++)	(+++)	(+++)	(+++)	(+++)	(+++)
$MarketTech\_RELATEDNESS$	(+)	(+)		(+)		(+)	(+)	(+)	(+)	(+)	(+)
$DesignTech\_RELATEDNESS$	(+)	(+)		(+)		(+)	(+)	(+)	(+)	(+)	(-)
					New design specializations						
$Design\_RELATEDNESS$	(+++)	(++)		(+++)		(++)	(++)	(++)	(++)	(++)	(++)
$TechDesign\_RELATEDNESS$	(++)	(+++)		(++)		(+++)	(+++)	(+++)	(+++)	(+++)	(+++)
$MarketDesign\_RELATEDNESS$	(+)	(+)		(+)		(+)	(+)	(+)	(+)	(+)	(+)
					New market specializations						
$Market\_RELATEDNESS$	(+++)	(+++)	(+++)	(+++)		(+++)	(+++)	(+++)	(+++)	(+++)	(+++)
$DesignMarket\_RELATEDNESS$	(+)	()	(+++)	()		()	()	()	()	()	()
$TechMarket\_RELATEDNESS$	(++)	(++)	(+++)	(+++)		(+++)	(+++)	(+++)	(++)	(+++)	(+++)

B: Baseline Results (Tables 3-5), P: Probit Results (Table A6), Add5: 5 countries added to the European sample (Table A7), I: Considering inventor location instead of applicant location (Table A8), F: Considering all filings (Table A9), C1: Fractional TMs 4-digit IPCs (Tables A10-A12), C2: Whole TMs 3-digit IPCs (Tables A10-A12), C3: Fractional TMs 3-digit IPCs (Tables A10-A12), M: Taking into account multiple IPC classifications (Table A13).

Table A6: Robustness check of Tables 3, 4 and 5, probit estimations. Marginal effects are displayed.

Panel A. New market specializations	(1)	(2)
VARIABLES	Europe	US
$Market\_RELATEDNESS_{r,i,t-1}$	0.242***	0.267***
	(0.011)	(0.013)
$TechMarket\_RELATEDNESS_{r,i,t-1}$	0.142***	0.311***
	(0.038)	(0.038)
$DesignMarket\_RELATEDNESS_{r,i,t-1}$	0.163***	0.011
	(0.040)	(0.044)
Observations	14,000	20,873
Panel B. New technology specializations	**(1)**	**(2)**
VARIABLES	Europe	US
$Tech\_RELATEDNESS_{r,i,t-1}$	0.127***	0.110***
	(0.002)	(0.002)
$MarketTech\_RELATEDNESS_{r,i,t-1}$	0.014***	0.017***
	(0.001)	(0.001)
$DesignTech\_RELATEDNESS_{r,i,t-1}$	0.034***	0.016***
	(0.002)	(0.002)
Observations	261,170	378,665
Panel C. New design specializations	**(1)**	**(2)**
VARIABLES	Europe	US
$Design\_RELATEDNESS_{r,i,t-1}$	0.110***	0.029***
	(0.004)	(0.003)
$MarketDesign\_RELATEDNESS_{r,i,t-1}$	0.010***	0.013***
	(0.003)	(0.002)
$TechDesign\_RELATEDNESS_{r,i,t-1}$	0.081***	0.083***
	(0.005)	(0.003)
Observations	88,873	120,203

Notes: Columns 1 and 2 of Panel A, B and C validate Columns 3 and 6 of Table 3, 4 and 5 respectively. All columns include region–period dummies. Standard errors are clustered at the region–class level and are displayed in parentheses.

Table A7: Robustness check of European case by adding EE, CY, LT, LV and MT.

Panel A. New market specializations	
VARIABLES	Europe
$Market\_RELATEDNESS_{r,i,t-1}$	0.223***
	(0.012)
$TechMarket\_RELATEDNESS_{r,i,t-1}$	0.134***
	(0.041)
$DesignMarket\_RELATEDNESS_{r,i,t-1}$	0.156***
	(0.042)
Constant	-0.006
	(0.073)
Observations	14,601
R-Squared	0.131
adj R-Squared	0.092
Panel B. New technology specializations	
VARIABLES	Europe
$Tech\_RELATEDNESS_{r,i,t-1}$	0.187***
	(0.003)
$MarketTech\_RELATEDNESS_{r,i,t-1}$	0.011***
	(0.001)
$DesignTech\_RELATEDNESS_{r,i,t-1}$	0.019***
	(0.003)
Constant	0.040***
	(0.008)
Observations	274,211
R-Squared	0.089
adj R-Squared	0.083
Panel C. New design specializations	
VARIABLES	Europe
$Design\_RELATEDNESS_{r,i,t-1}$	0.084***
	(0.004)
$MarketDesign\_RELATEDNESS_{r,i,t-1}$	0.014***
	(0.002)
$TechDesign\_RELATEDNESS_{r,i,t-1}$	0.051***
	(0.006)
Constant	0.037*
	(0.020)
Observations	93,482
R-Squared	0.082
adj R-Squared	0.073

Note: The dependent variable in all regressions is $Entry_{r,i,t}$. All regressions are estimated via ordinary least squares (OLS). All columns include region–period and class-period dummies. Standard errors are clustered at the region–class level and are displayed in parentheses. Panels A, B and C validate Column 3 of Tables 3, 4 and 5 respectively.

Table A8: Robustness check of Tables 3, 4 and 5, considering inventor location instead of applicant location.

Panel A. New market specializations	(1)	(2)
VARIABLES	Europe	US
$Market\_RELATEDNESS_{r,i,t-1}$	0.213***	0.230***
	(0.012)	(0.015)
$TechMarket\_RELATEDNESS_{r,i,t-1}$	0.281***	0.226***
	(0.050)	(0.038)
$DesignMarket\_RELATEDNESS_{r,i,t-1}$	0.129***	-0.041
	(0.041)	(0.043)
Constant	-0.367***	-0.104
	(0.077)	(0.064)
Observations	14,336	21,011
R-Squared	0.131	0.091
adj R-Squared	0.093	0.054
Panel B. New technology specializations	**(1)**	**(2)**
VARIABLES	Europe	US
$Tech\_RELATEDNESS_{r,i,t-1}$	0.245***	0.225***
	(0.004)	(0.003)
$MarketTech\_RELATEDNESS_{r,i,t-1}$	0.016***	0.015***
	(0.002)	(0.001)
$DesignTech\_RELATEDNESS_{r,i,t-1}$	0.015***	0.011***
	(0.003)	(0.003)
Constant	0.011	0.214***
	(0.010)	(0.009)
Observations	261,680	374,802
R-Squared	0.092	0.125
adj R-Squared	0.086	0.121
Panel C. New design specializations	**(1)**	**(2)**
VARIABLES	Europe	US
$Design\_RELATEDNESS_{r,i,t-1}$	0.082***	0.061***
	(0.004)	(0.004)
$MarketDesign\_RELATEDNESS_{r,i,t-1}$	0.015***	0.004***
	(0.002)	(0.002)
$TechDesign\_RELATEDNESS_{r,i,t-1}$	0.066***	0.090***
	(0.006)	(0.005)
Constant	0.032	0.061***
	(0.020)	(0.018)
Observations	91,686	123,493
R-Squared	0.082	0.113
adj R-Squared	0.073	0.105

Notes: Columns 1 and 2 of Panel A validate Columns 3 and 6 of Table 3, 4 and 5 respectively. All columns include region–period and class-period dummies. Standard errors are clustered at the region–class level and are displayed in parentheses.

Table A9: Robustness check of Tables 3, 4 and 5, considering all filings.

Panel A. New market specializations	(1)	(2)
VARIABLES	Europe	US
$Market\_RELATEDNESS_{r,i,t-1}$	0.198***	0.227***
	(0.012)	(0.015)
$TechMarket\_RELATEDNESS_{r,i,t-1}$	0.097**	0.128***
	(0.038)	(0.049)
$DesignMarket\_RELATEDNESS_{r,i,t-1}$	0.184***	0.022
	(0.042)	(0.046)
Constant	0.001	0.138**
	(0.074)	(0.059)
Observations	14,322	20,545
R-Squared	0.123	0.101
adj R-Squared	0.084	0.064
Panel B. New technology specializations	**(1)**	**(2)**
VARIABLES	Europe	US
$Tech\_RELATEDNESS_{r,i,t-1}$	0.205***	0.180***
	(0.004)	(0.003)
$MarketTech\_RELATEDNESS_{r,i,t-1}$	0.016***	0.005***
	(0.001)	(0.001)
$DesignTech\_RELATEDNESS_{r,i,t-1}$	0.026***	0.012***
	(0.003)	(0.003)
Constant	0.091***	-0.054***
	(0.009)	(0.012)
Observations	265,534	379,247
R-Squared	0.091	0.115
adj R-Squared	0.085	0.110
Panel C. New design specializations	**(1)**	**(2)**
VARIABLES	Europe	US
$Design\_RELATEDNESS_{r,i,t-1}$	0.084***	0.056***
	(0.005)	(0.004)
$MarketDesign\_RELATEDNESS_{r,i,t-1}$	0.015***	0.006***
	(0.002)	(0.002)
$TechDesign\_RELATEDNESS_{r,i,t-1}$	0.049***	0.098***
	(0.006)	(0.005)
Constant	0.005	0.065***
	(0.021)	(0.018)
Observations	91,686	123,493
R-Squared	0.082	0.114
adj R-Squared	0.073	0.106

Notes: Columns 1 and 2 of Panel A, B, C validate Columns 3 and 6 of Table 3, 4 and 5. All models include region–period and class-period dummies. Standard errors are clustered at the region–class level and are displayed in parentheses.

Table A10: Robustness checks with alternative class counting, models for new market specializations.

	Europe			US		
	Fractional TMs 4-digit IPCs	Whole TMs 3-digit IPCs	Fractional TMs 3-digit IPCs	Fractional TMs 4-digit IPCs	Whole TMs 3-digit IPCs	Fractional TMs 3-digit IPCs
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
$Market\_RELATEDNESS_{r,i,t-1}$	0.216***	0.221***	0.215***	0.204***	0.237***	0.198***
	(0.013)	(0.012)	(0.012)	(0.015)	(0.015)	(0.015)
$TechMarket\_RELATEDNESS_{r,i,t-1}$	0.094**	0.092***	0.081***	0.188***	0.123***	0.172***
	(0.037)	(0.033)	(0.029)	(0.042)	(0.029)	(0.028)
$DesignMarket\_RELATEDNESS_{r,i,t-1}$	0.173***	0.154***	0.173***	0.021	-0.016	0.025
	(0.037)	(0.041)	(0.037)	(0.041)	(0.042)	(0.040)
Constant	0.037	0.035	0.063	0.090	0.019	-0.103
	(0.067)	(0.073)	(0.069)	(0.060)	(0.067)	(0.067)
Observations	15,070	14,336	15,070	21,381	21,011	21,381
R-Squared	0.116	0.130	0.116	0.089	0.090	0.090
adj R-Squared	0.078	0.091	0.079	0.053	0.054	0.054

Note: The dependent variable in all regressions is $Entry_{r,i,t}$. All regressions are estimated via ordinary least squares (OLS). All columns include region–period and Nice class-period dummies. Standard errors are clustered at the region–class level and are displayed in parentheses.

Table A11: Robustness checks with alternative class counting, models for new technology specializations.

	Europe			US		
	Fractional TMs 4-digit IPCs	Whole TMs 3-digit IPCs	Fractional TMs 3-digit IPCs	Fractional TMs 4-digit IPCs	Whole TMs 3-digit IPCs	Fractional TMs 3-digit IPCs
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
$Tech\_RELATEDNESS_{r,i,t-1}$	0.187***	0.169***	0.166***	0.181***	0.252***	0.253***
	(0.003)	(0.009)	(0.009)	(0.003)	(0.008)	(0.008)
$MarketTech\_RELATEDNESS_{r,i,t-1}$	0.012***	0.040***	0.044***	0.004***	0.030***	0.029***
	(0.001)	(0.005)	(0.005)	(0.001)	(0.004)	(0.004)
$DesignTech\_RELATEDNESS_{r,i,t-1}$	0.019***	0.069***	0.068***	0.011***	0.047***	0.047***
	(0.003)	(0.009)	(0.009)	(0.003)	(0.008)	(0.008)
Constant	0.049***	-0.075**	-0.062**	-0.060***	-0.099***	-0.105***
	(0.008)	(0.031)	(0.031)	(0.012)	(0.032)	(0.032)
Observations	268,815	50,698	50,698	379,247	69,199	69,199
R-squared	0.089	0.084	0.084	0.115	0.124	0.124
adj R-Squared	0.083	0.070	0.070	0.110	0.112	0.112

Note: The dependent variable in all regressions is $Entry_{r,i,t}$. All regressions are estimated via ordinary least squares (OLS). All columns include region–period and IPC class-period dummies. Standard errors are clustered at the region–class level and are displayed in parentheses.

Table A12: Robustness checks with alternative class counting, models for new design specializations.

	Europe			US		
VARIABLES	Fractional TMs 4-digit IPCs (1)	Whole TMs 3-digit IPCs (2)	Fractional TMs 3-digit IPCs (3)	Fractional TMs 4-digit IPCs (4)	Whole TMs 3-digit IPCs (5)	Fractional TMs 3-digit IPCs (6)
$Design\_RELATEDNESS_{r,i,t-1}$	0.083***	0.090***	0.090***	0.056***	0.081***	0.081***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
$MarketDesign\_RELATEDNESS_{r,i,t-1}$	0.018***	0.014***	0.018***	0.005***	0.006***	0.005***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$TechDesign\_RELATEDNESS_{r,i,t-1}$	0.048***	0.035***	0.034***	0.098***	0.051***	0.051***
	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Constant	0.048**	0.014	0.017	0.062***	0.058***	0.055***
	(0.020)	(0.021)	(0.021)	(0.018)	(0.019)	(0.019)
Observations	91,686	91,686	91,686	123,493	123,493	123,493
R-squared	0.082	0.082	0.082	0.114	0.111	0.111
adj R-Squared	0.0726	0.0722	0.0723	0.106	0.103	0.103

Note: The dependent variable in all regressions is $Entry_{r,i,t}$. All regressions are estimated via ordinary least squares (OLS). All columns include region–period and Locarno class-period dummies. Standard errors are clustered at the region–class level and are displayed in parentheses.

Table A13: Robustness check of Tables 3, 4 and 5, taking into account multiple IPC classifications.

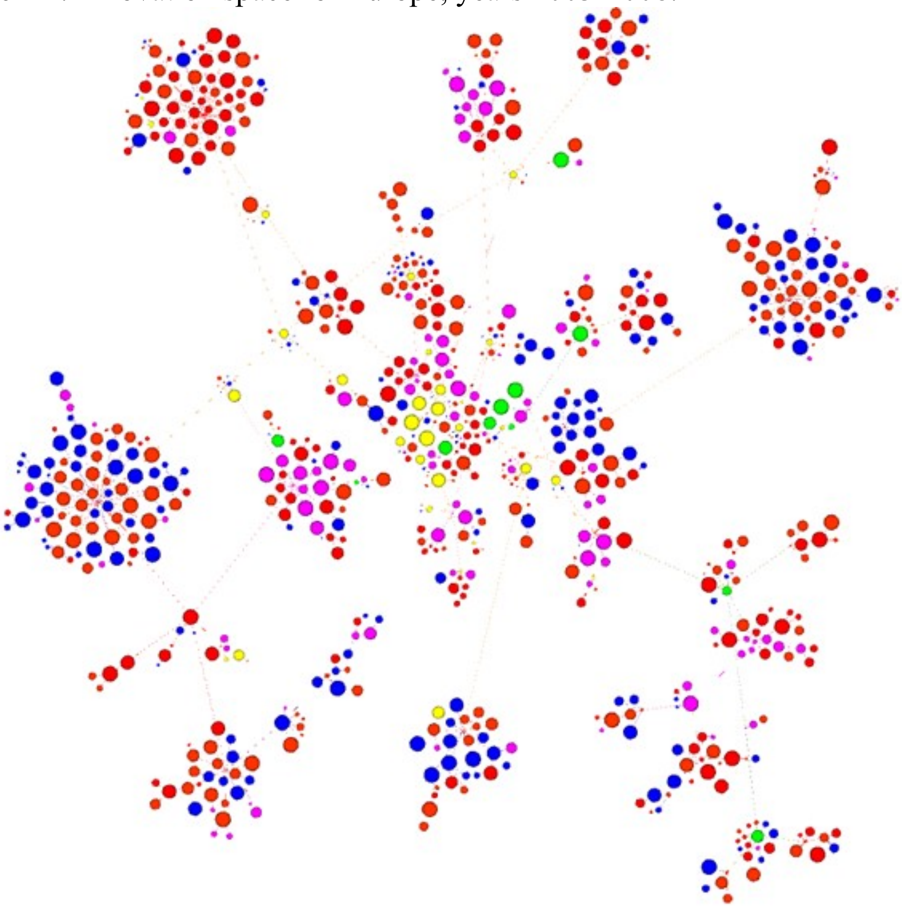
Panel A. New trademark specializations	(1)	(2)
VARIABLES	Europe	US
$Market\_RELATEDNESS_{r,i,t-1}$	0.217***	0.240***
	(0.012)	(0.015)
$TechMarket\_RELATEDNESS_{r,i,t-1}$	0.157***	0.152***
	(0.041)	(0.042)
$DesignMarket\_RELATEDNESS_{r,i,t-1}$	0.144***	-0.022
	(0.041)	(0.043)
Constant	0.031	-0.071
	(0.071)	(0.073)
Observations	14,336	21,011
R-Squared	0.130	0.090
adj R-Squared	0.0914	0.0532

Panel B. New technology specializations	(1)	(2)
VARIABLES	Europe	US
$Tech\_RELATEDNESS_{r,i,t-1}$	0.227***	0.249***
	(0.004)	(0.004)
$MarketTech\_RELATEDNESS_{r,i,t-1}$	0.014***	0.016***
	(0.002)	(0.001)
$DesignTech\_RELATEDNESS_{r,i,t-1}$	0.027***	-0.008**
	(0.003)	(0.003)
Constant	0.065***	-0.207***
	(0.009)	(0.014)
Observations	261,692	373,078
R-Squared	0.085	0.110
adj R-Squared	0.0795	0.105

Panel C. New design specializations	(1)	(2)
VARIABLES	Europe	US
$Design\_RELATEDNESS_{r,i,t-1}$	0.084***	0.064***
	(0.004)	(0.004)
$MarketDesign\_RELATEDNESS_{r,i,t-1}$	0.013***	0.005***
	(0.002)	(0.002)
$TechDesign\_RELATEDNESS_{r,i,t-1}$	0.058***	0.100***
	(0.006)	(0.006)
Constant	0.003	0.062***
	(0.020)	(0.018)
Observations	91,686	123,493
R-Squared	0.082	0.113
adj R-Squared	0.0726	0.105

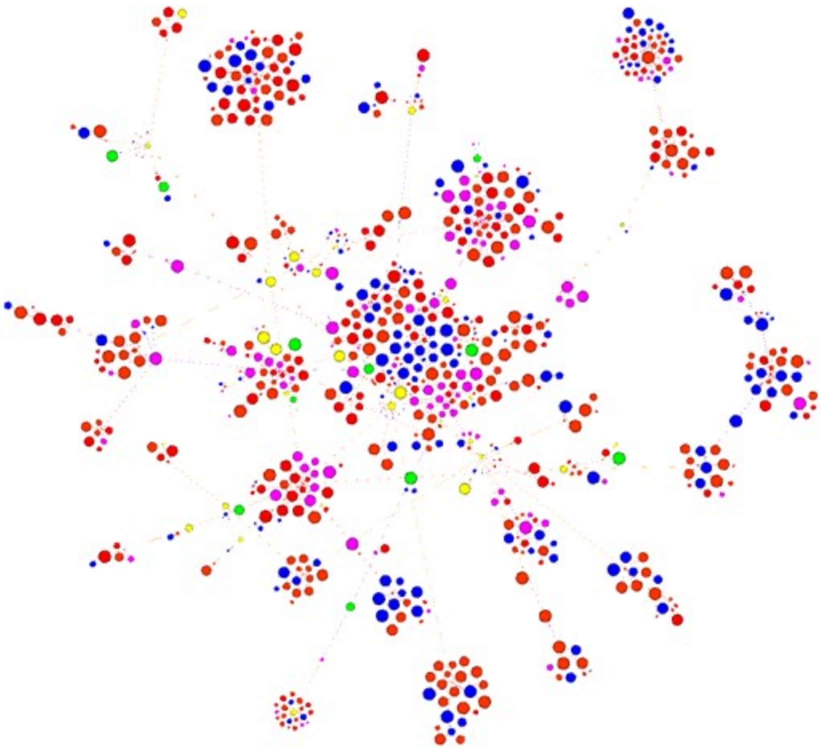
Notes: Columns 1 and 2 of Panel A, B and C validate Columns 3 and 6 of Table 3, 4 and 5 respectively. All columns include region–period dummies. Standard errors are clustered at the region–class level and are displayed in parentheses.

Figure A1: Innovation space for Europe, years 2003-2008.



Technology IPC classes: A, D
Technology IPC classes: B, C
Technology IPC classes: E, F, G, H
Design Locarno classes
Market Nice product classes: 1-34
Market Nice service classes: 35-45

Figure A2: Innovation space for US, years 2003-2008.



Technology IPC classes: A, D

Technology IPC classes: B, C

Technology IPC classes: E, F, G, H

Design Locarno classes

Market Nice product classes: 1-34

Market Nice service classes: 35-45