Economic
Research
Working Paper
No. 80/2024

Global Trends in Innovation Patterns
A Complexity Approach

Ricardo Hausmann, Muhammed A. Yildirim, Christian Chacua, Matte Hartog, Shreyas Gadgin Matha

WIPO

Global Trends in Innovation Patterns: A Complexity Approach

Ricardo Hausmann* Muhammed A. Yildirim*†. Christian Chacua*
Matte Hartog* Shreyas Gadgin Matha*

April 2024

Abstract

Technological know-how in a country shapes its growth potential and competitiveness. Scientific publications, patents, and international trade data offer complementary insights into how ideas from science, technology, and production evolve, combine, and are transformed into capabilities. Analyzing their trajectories enables a more comprehensive and multifaceted understanding of the whole innovation process, from generating ideas to internationally commercializing products. We analyze the production patterns in these three domains, documenting the differences between advanced and emerging market economies. We find that future income, patenting, and publishing growth correlate with the economic complexity indices calculated from these domains. Capabilities embedded in the country also shape future diversification opportunities and make the innovation process path dependent. Lastly, we also show that diversification opportunities can be inferred across innovation domains.

Keywords: economic complexity, innovation complexity, scientific complexity.

JEL Codes: O25, O30, O38, F60.

Disclaimer: The views expressed in this article are those of the authors and do not necessarily reflect the views of the World Intellectual Property Organization (WIPO) or its Member States.

Acknowledgments: We are indebted to Andrea Morrison, Sebastian Bustos, Julio Raffo, Federico Moscatelli, Paolo Aversa, Prince Oguguo, Basil Mahfouz, and Hakan Özalp for their comments and suggestions. Any mistake remains ours.

*The Growth Lab, Center for International Development - Harvard University, Cambridge, MA, United States.
†Department of Economics, College of Administrative Sciences and Economics - Koç University, Istanbul, Turkey.

1 Economic complexity approach

Technological progress is a cornerstone for economic growth and competitiveness. As we argued in the accompanying paper (Hausmann et al., 2024), technology does not diffuse adequately in space, which explains income differences among countries. Across nations, regions, and industries, the dynamics of technological progress are shaped not only by the number of new ideas generated but also by the sophistication and interconnectedness of ideas within the economic system. We argue that the interplay between ideas generated from multiple cognitive domains related to economic production influences an economy's development trajectory. Scientific publications, patents, and international trade data offer complementary insights into how ideas from science, technology, and production evolve, combine, and are transformed into capabilities by countries. Analyzing their trajectories enables a more comprehensive and multifaceted understanding of the whole innovation process, from generating ideas to internationally commercializing products.

This paper introduces and applies economic complexity metrics to describe global trends of innovation complexity. It extends concepts from the economic complexity framework on trade to the analysis of scientific and technological progress, measured through scientific publications and patents. Following the Scrabble analogy presented in Hausmann et al. (2024), we discuss measures of letters used in three domains (scientific, technological, and industrial capabilities) and provide some estimations of these metrics for the period 2000-2020. Our computations rely on country-level data on scientific publications extracted from OpenAlex (Priem et al., 2022), patents compiled by WIPO, and international trade from the United Nations COMTRADE dataset, processed by Bustos & Yildirim (2023).

Our analyses cover four different sets of measures of capabilities. First, we present results that build on the production levels and the concept of Revealed Comparative Advantage (RCA) (Balassa, 1965). The RCA provides information to identify broad trends in the production and geographical distribution of scientific, technological, and industrial outcomes. Second, we cover metrics that capture an economy's degree of sophistication and knowledge accumulation. They rely on generalized versions of the Economic Complexity Index (Hausmann et al., 2014; Hidalgo & Hausmann, 2009) and are used to examine the relationship between complexity and economic growth. Third, we introduce measures that exhibit the temporal evolution of capabilities and the process of scientific and technological diversification. Those measures capture the overlap between the letters (i.e., Relatedness), generalizing the product space methodology (Hidalgo et al., 2007) developed for products in international trade to scientific fields and technological classes (Balland & Boschma, 2022; Petralia et al., 2017). Finally, we quantify the interplay among scientific, technological, and industrial capabilities. They are used to explore the potential for countries to achieve complex technologies based on their scientific and industrial capabilities. In Box 1, we briefly explain the particular metrics used in this paper.

We document several patterns in global innovative activity. First, we show that scientific publishing and patenting are far more concentrated in a few countries than exports. Countries with a high diversity of innovative activity (as measured by patents, publications, or exports) tend to produce more unique (less ubiquitous) innovations. Next, we compute clusters of countries based on the type of ideas they produce. We identify two clusters that differ significantly based on their research spending and GDP per capita but not on their population. This indicates that there might be differences between the types of scientific and technological activity conducted by advanced economies and emerging market economies. Taking this further, we find that the sophistication of a country's innovative activity (as measured by Economic Complexity Indices of patents, publications, or exports) positively correlates with growth in GDP per capita, patenting, scientific publications, and exports. Next, we show that countries exhibit related diversification in innovation. That is, they are more likely to diversify their innovation portfolio by moving into nearby export products in the product space, nearby scientific fields in the scientific field space, and nearby patent classes in the technology class space. Lastly, we find this pattern of related diversification across innovation domains. For example, a country that exports relatively much of certain products will likely grow its patenting activity in related technology classes.

The remainder of this paper is organized as follows. Section 2 describes the data used in the analysis. Sections 3 to 6 analyze the four sets of metrics that guide our analyses and their relevance for innovation policy discussions. Section 7 concludes with a discussion on policy implications. Additionally, we include three appendix sections that cover the details of the metrics' mathematical formulations (A), the empirical data methods (B), and supplementary results (C).

Box 1: Main concepts

Here, we introduce various concepts frequently used in the study of economic complexity. We present our definitions and interpretations within the context of international trade, while also highlighting their applicability to technology and scientific outputs.

- **Revealed Comparative Advantage (RCA)** is a metric that determines a country's relative advantage or disadvantage in producing a specific product compared to other nations. Essentially, if a country has an RCA greater than one for a product, that country has a comparative advantage in producing that product.

- Two other fundamental concepts, **diversity** and **ubiquity**, intertwine closely with RCA. While diversity assesses the range of products a country produces with a comparative advantage, ubiquity evaluates how widespread a product is among all countries. Both offer insights into a nation's productive capabilities and global demand.

- The **Economic Complexity Index (ECI)** measures the sophistication and knowledge accumulation of an economy. It captures an economy's ability to produce a wide variety of complex products. Higher values of ECI signify that an economy produces diverse products that are less commonly manufactured globally, revealing depth in knowledge and capabilities.

- **Relatedness** examines the shared capabilities, skills, and know-how necessary for producing two different products. Two products are considered more related if countries that competitively produce one product are also frequently competitive in the other.

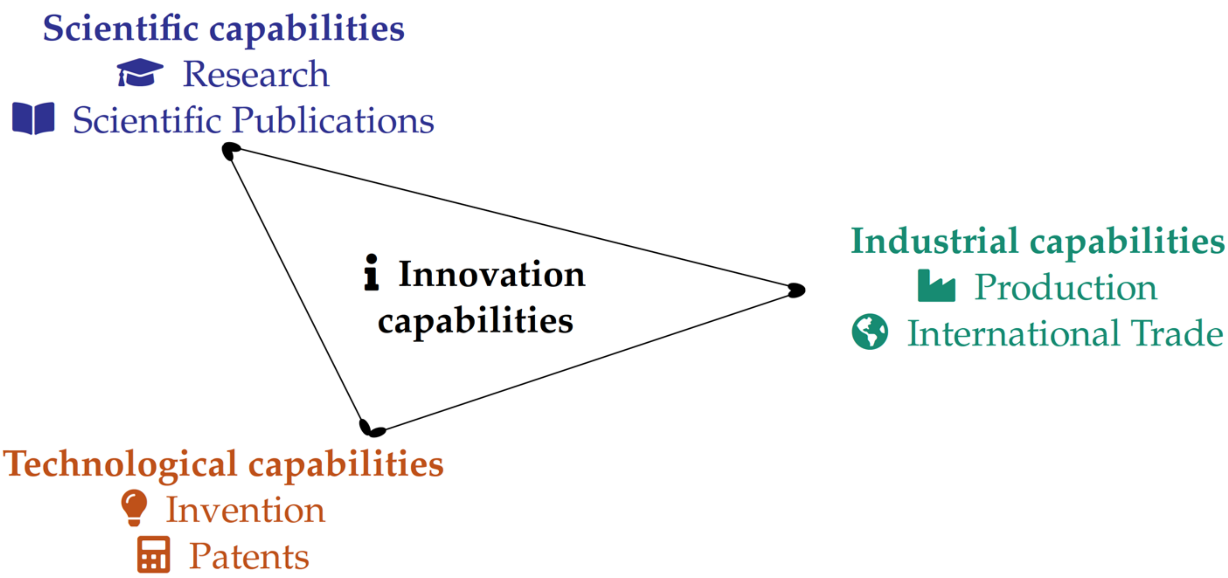
- The relatedness of products can be visualized on a network called the **Product Space**. In the Product Space, products are represented as nodes, and the proximity between products indicates how often they co-occur in countries' export baskets, suggesting shared capabilities in production.

Besides international trade and products, we apply these concepts to scientific and technological ideas. We create scientific field spaces based on data from scientific publications and technology spaces based on patenting data. We explore various definitions of relatedness. Further, we quantify the sophistication involved in innovating in various scientific fields and technologies and chart paths of least resistance for countries to move towards more complex scientific fields and technologies.

2 Data

Our multidimensional approach builds on three primary sources of information that offer a comprehensive view of the innovation process: scientific publications, patents, and international trade (Figure 1). Scientific publications capture the creation of knowledge that could translate into scientific capabilities. Patents unveil the inventions that could translate into technological capabilities. International trade data reveal the current industrial capabilities of an economy. Together, these three measures provide a more extensive and complementary view of the complex nature of the innovation process, based on the capabilities present in a location.

Figure 1: A multidimensional view of innovation capabilities



Source: Own construction. Our multidimensional view of innovation capabilities considers three types of capabilities: scientific, technological, and industrial capabilities. Hence, we analyze data on three types of activities: research, as measured in scientific publications; invention, as measured in patents; and industrial production, as measured in trade data.

Scientific publications offer insights into the ideas that mainly originated from scientific research. It captures the nascent ideas from academia and basic research institutions that might underpin future productive innovations. Although not all scientific knowledge leads to productive innovations, as it is not its primary goal, a significant share of productive innovations originated from basic science (Mazzucato, 2015; Gruber & Johnson, 2019). We rely on data from OpenAlex (Priem et al., 2022) in all our computations related to scientific publications. Although most bibliometric databases, including OpenAlex, claim to have global coverage of scientific publications, we are aware that they do not contain representative data for all countries. Hence, the global trends presented here may reflect biases in coverage, which may disproportionally affect countries in the global south.[1]

Patents are an indicator of invention, one of the intermediate steps of the innovation process. It assesses the potential transformation of ideas into market products. An economy that consistently generates patents in a sector likely has productive know-how and capabilities in that sector. Even though we recognize the limitations of using patent data, it remains a valuable source for gauging innovative trajectories.[2] Our patent data has been compiled by WIPO, combining data from multiple sources, primarily EPO's Patstat 2023 and WIPO's PatentScope.

Trade datasets, from which we extract export data, serve as a benchmark and primarily show the current industrial capabilities of an economy. This information reveals what is feasible for a nation to produce and where it stands in the global economic landscape. Our trade data comes mainly from international data collected by UN COMTRADE from customs offices and further cleaned by Bustos & Yildirim (2023). This cleaning procedure tries to account for differences in data reported by exporters and importers, as well as the quality of data reporting by various countries.

In the three considered datasets, we analyze data at the country level for the period 2000-2020. We focus on countries to describe global trends, but we acknowledge that the design of particular innovation policies requires analysis at more disaggregated levels. Moreover,

[1] Most bibliometric databases, including OpenAlex, still have poor coverage of non-English documents, local journals, and journals that are only available in print (Ansorge, 2023). As those characteristics are not randomly distributed across countries (e.g., non-English speaking countries may disproportionally write non-English articles), scientific publication data may not be fully representative of the geographic distribution of scientific knowledge.

[2] As the use of patents varies by several factors (e.g., industry, firm size, type of innovation, among others) (Mezzanotti & Simcoe, 2023; Cohen et al., 2000; Levin et al., 1987; Harabi, 1995), the statistical inference based on patent data is limited. In addition, patenting practices can differ across countries and may respond to strategic behaviors (Lemley & Shapiro, 2007; Golden, 2007; Henkel, 2022).

our study period is not large enough to understand the dynamics of all the stages of innovation processes, which in some cases can span multiple decades and require a more detailed assessment of how individual ideas are transformed into final products. However, it allows us to assess the current state of scientific, technological, and industrial capabilities, as well as provide insights into their geographical distribution, degree of sophistication, recent evolution, and potential connections.

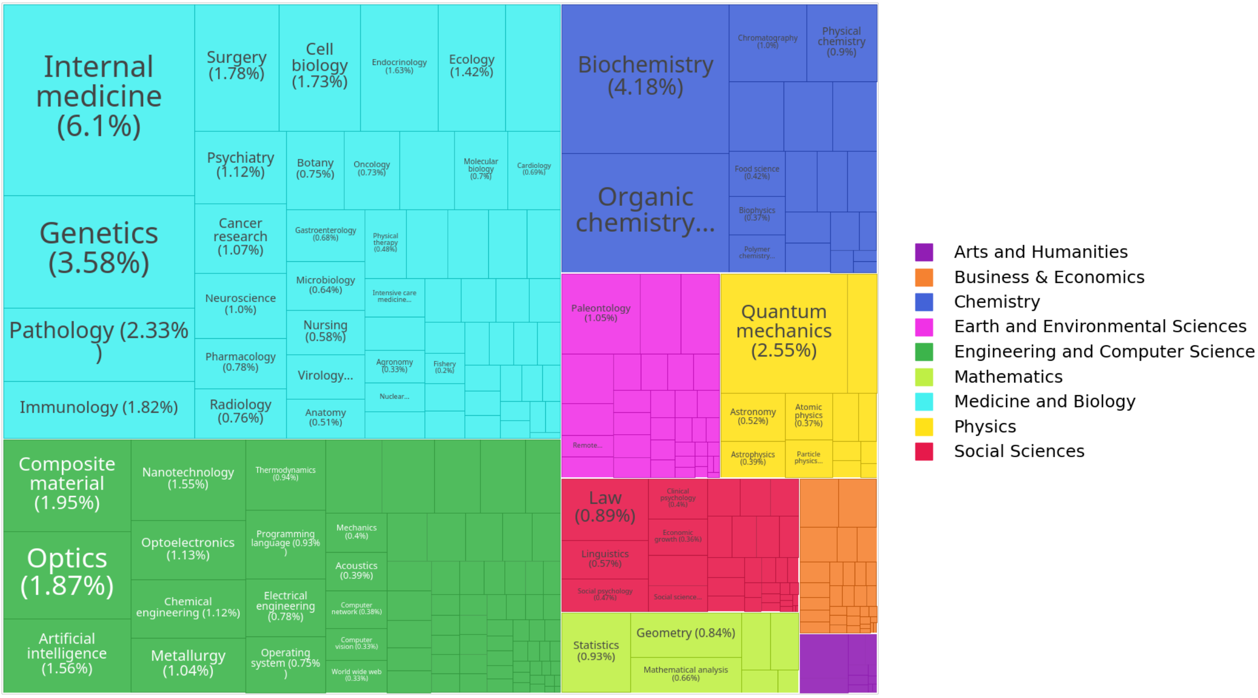
As we focus on measuring the capabilities of each country, we assign scientific publications, patents, and exports to the places where they are produced. For scientific publications, we assign papers to countries based on the location of their authors' institutional affiliations. For patents, we rely on the location of the listed inventors in a patent family. For exports, we use the exporter's location from a cleaned version of UN COMTRADE data (Bustos & Yildirim, 2023). When measuring the number of scientific publications and patents in a country, we compute fractional counts based on the number of distinct countries, not on the number of different authors or inventors. Furthermore, to make patent data internationally comparable, when we refer to patent counts, we count international patent families (Miguelez et al., 2019), and not individual patent applications. In addition, we apply a series of filters to remove countries for which the data on a particular dimension is not meaningful for statistical analysis. More details on the specific data sources and methods can be found in the appendix B.

3 Global trends

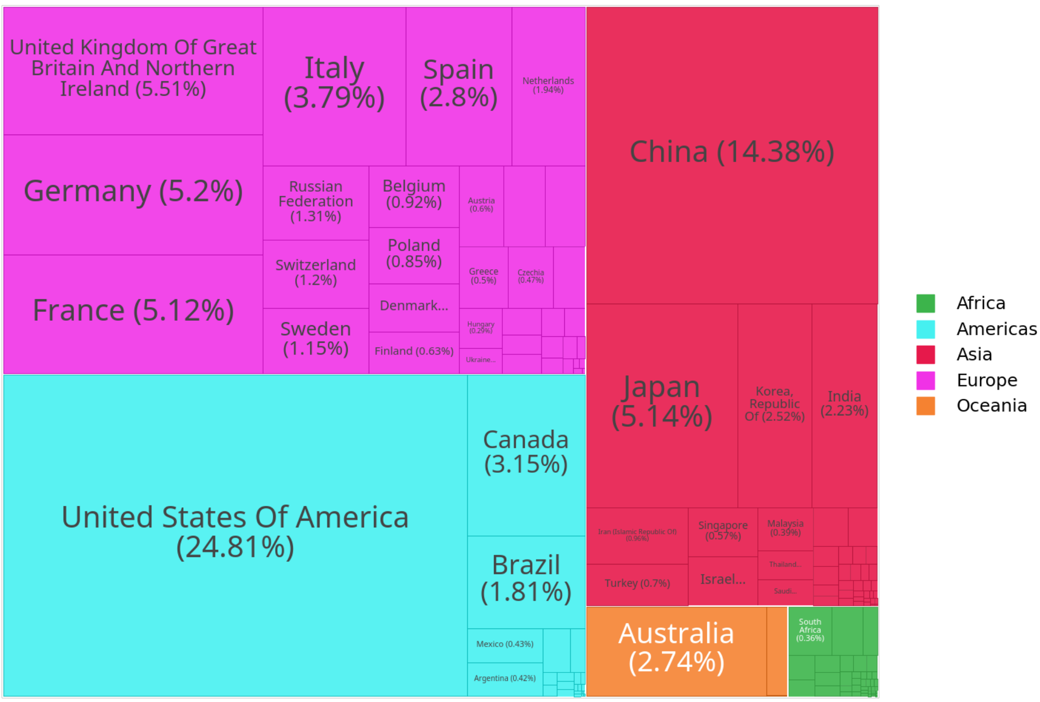
As illustrated in Figure 2a, between 2000 and 2020, most publications worldwide were in the fields of Medicine and Biology, followed by Engineering and Computer Science. Both primary academic fields account for more than half of all scientific publications. Afterward, we find the fields of Chemistry, Earth and Environmental Sciences, and Physics. The social sciences' share is relatively small at about 15%. By country of authors' affiliation (Figure 2b), the United States accounts for around 25% of all publications, preceded by China at about 14%. The following four countries, except for Japan, are located in Europe. Africa, as a whole continent, only accounts for a small fraction, less than 2%, of the global scientific publications in our data.

Figure 3a shows the distribution of patents by technological class. The major technology groups of Physics, Electricity, Performing Operations, and Human Necessities, each account for about 15% of all patents. The largest technologies at the IPC4 level are related to electric digital data processing and semiconductors. The United States, similarly to publications, accounts for about a quarter of all patents, by country of inventors' location (Figure 3b). The following main countries of invention are located in Asia, with Japan, Korea, and China contributing together to above 40% of global patenting. In Europe, Germany and France stand out with 9% and 3% of world patenting, respectively.

Figure 2: Scientific publications worldwide between 2000 and 2020



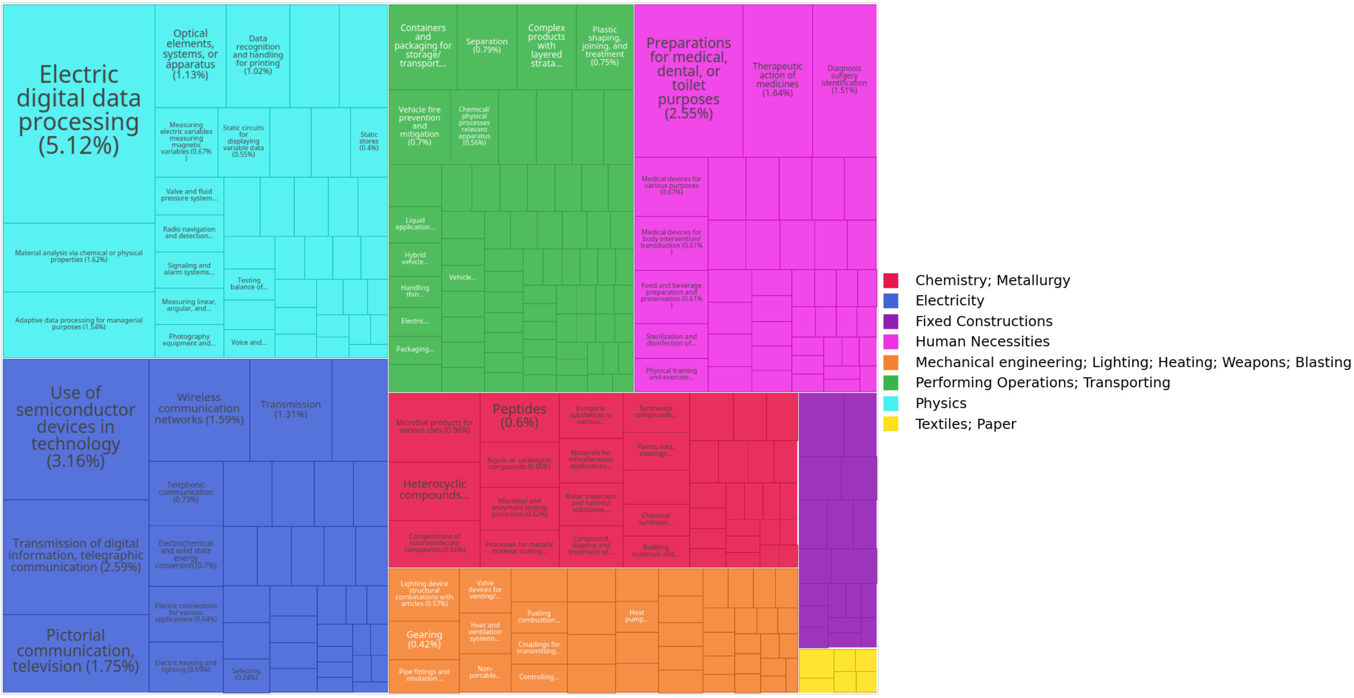
(a) By field



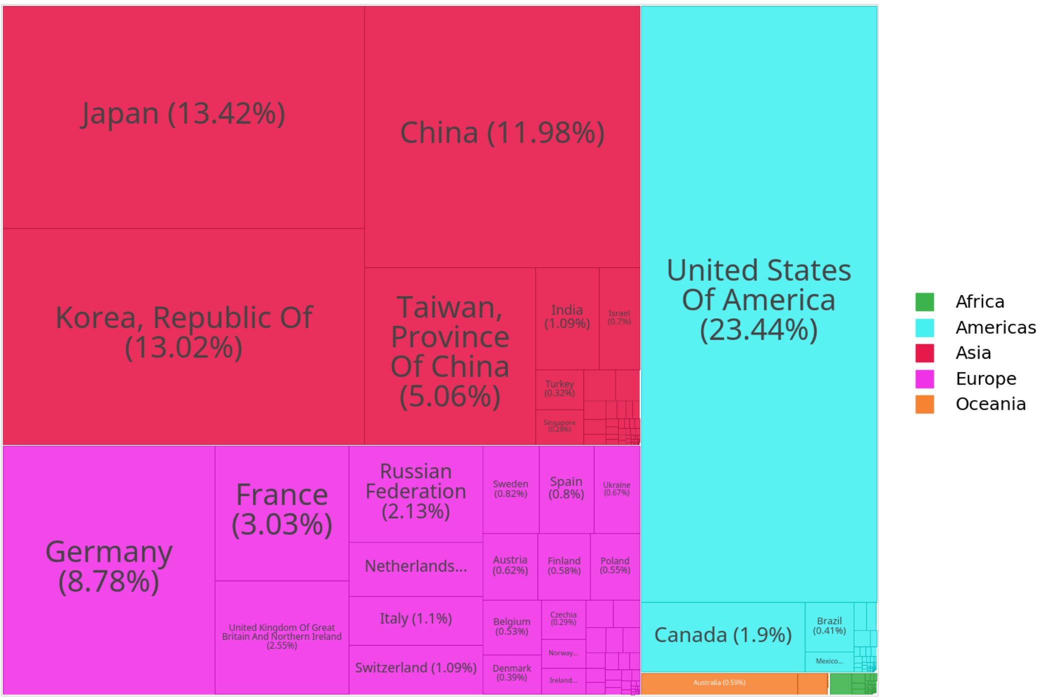
(b) By country

Source: Own construction, using scientific publication data from OpenAlex.

Figure 3: Patents worldwide between 2000 and 2020



(a) By technology



(b) By country

Source: Own construction, using patent data compiled by WIPO.

Interestingly, the creation of scientific and technological knowledge is more concentrated than the production of traded goods. As shown in Figures 4a and 4b, the per capita production of patents and scientific publications is more concentrated than the per capita production of exports or even dollars of income. This is more evident at the upper tail of the distributions of GDP, patents, and publications, as the top 0.01% countries account for around 25% of their total global production. In contrast, the top 0.01% exporter countries only account for about 11% of the world's exports. When moving down the distributions, patents, and scientific publications are even more concentrated than income. The top 1% producers of patents account for 44% of the total patents. For scientific publications, this value is near 39%, while for income, it is about 42%. Again, exports are the least concentrated of the considered dimensions, as the top 1% exporters account for around 20% of the world's exports. Despite the data limitations associated with patents and scientific publications, these figures reflect the nature of scientific and technological knowledge production: it is highly unequal and concentrated in a few top producers.[3] Observing such dramatic concentration in the production of scientific and technological knowledge is counter-intuitive, given that we expect ideas to flow more seamlessly than tangible products. Knowledge is non-rivalrous and non-excludable in consumption and is thus considered a global public good (Nelson, 1959; Arrow, 1962; Stiglitz, 1999). One would then expect that it should be relatively easy to build on freely accessible existing knowledge compared to expanding trade and that knowledge production should be more evenly distributed across countries. The observed concentration suggests significant barriers or advantages in knowledge production, contradicting the notion of knowledge as a freely diffusing public good (Breschi & Lissoni, 2001; Cowan et al., 2000).

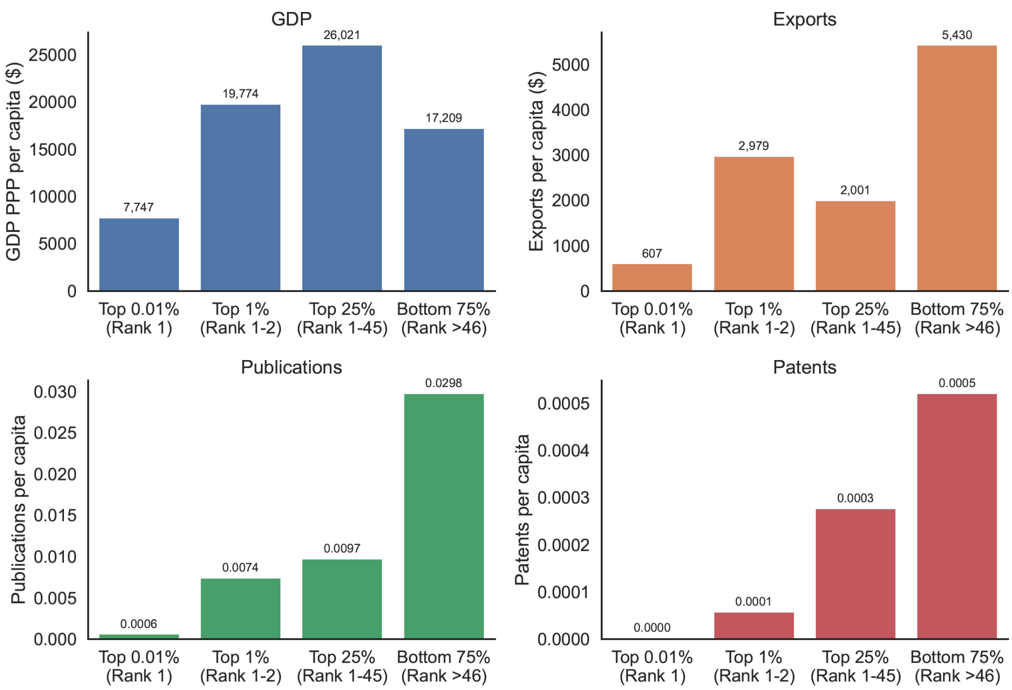
Scientific and technological knowledge production is disproportionately concentrated in high-income countries. As shown in panel 4c, the production of scientific publications patents,

[3] Alternatively, one could argue that the large concentration of scientific publications and patents only reflects the lack of coverage beyond the top countries or that they are not suitable to perform worldwide comparisons.

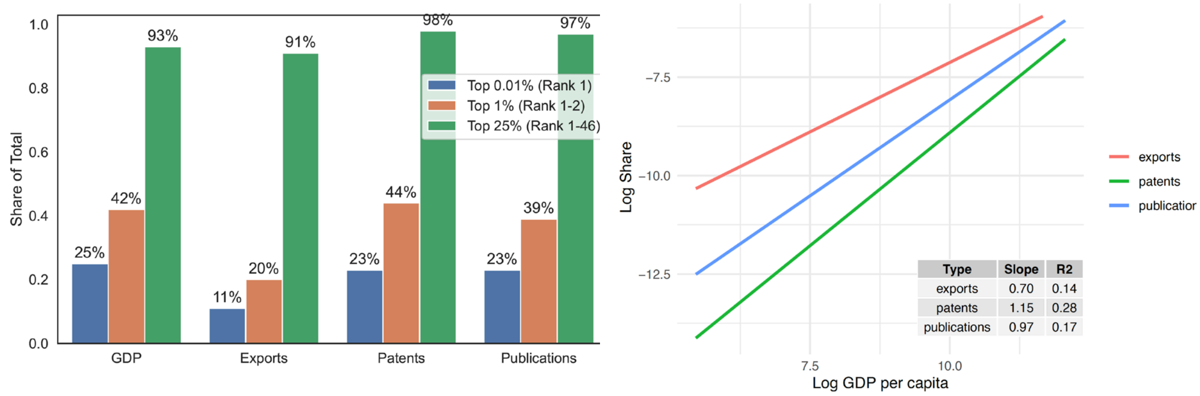
Although no dataset fully covers all the scientific publications and patents produced worldwide, our main sources of information and data processing methods try to handle potential coverage biases (see details of the data methods in Appendix B). Certainly, patents and scientific publications are not the only way of producing scientific and technological knowledge. However, they give an idea of the direction of science and innovation, are recorded more systematically, and are available for multiple countries and years.

and exports positively correlates with the per capita income level. More significantly, changes in the production of scientific publications and patents seem to respond proportionally to changes in the GDP per capita. The sensitivity (elasticity) of per capita income to output concentration (slope coefficient) is around one for both patents (i.e., 1.15) and scientific publications (i.e., 0.97). This behavior differs from the pattern found in export data, where the same sensitivity measure is close to 0.70. Without controlling for additional confounding factors or providing robust estimations, our preliminary results suggest that the sensitivity of patent production to per capita income could be 60% higher than that of exports and around 20% higher than that of scientific publications.

Figure 4: Inequality in scientific publications, patents, and exports, 2000-2020.



(a) Per Capita production



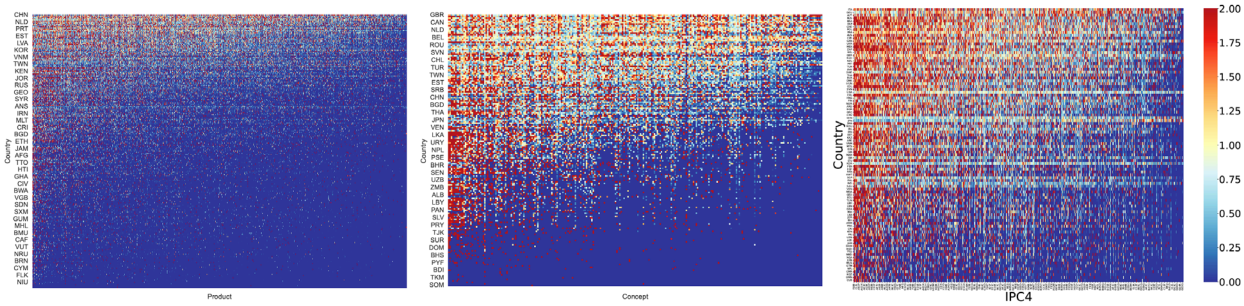
(b)Shares in total production

(c) Share in total production vs. GDP per capita

Source: Own construction, using export data from UN COMTRADE, scientific publications from OpenAlex, patents compiled by WIPO, and income data from the World Bank. Note: We only consider countries with at least 100,000 people. In panel C.4a, the values correspond to the per-capita values of each group (adding all the outcomes and populations of the countries listed in a given category).

Figure 5 shows the extent to which countries specialize in products (exports), technologies, and scientific publications. This is measured by the Revealed Comparative Advantage, or RCA (Balassa, 1965), which compares the share of a country's activity in one of the areas to the area's global activity, dividing one by the other, to obtain a degree of over- or under- representation (Section A elaborates on the technical details). Countries are distinctly more specialized in scientific output and technologies than they are in products, suggesting the former require more specific or deeper capabilities than the latter.

Figure 5: RCA Matrices, 2000 – 2020



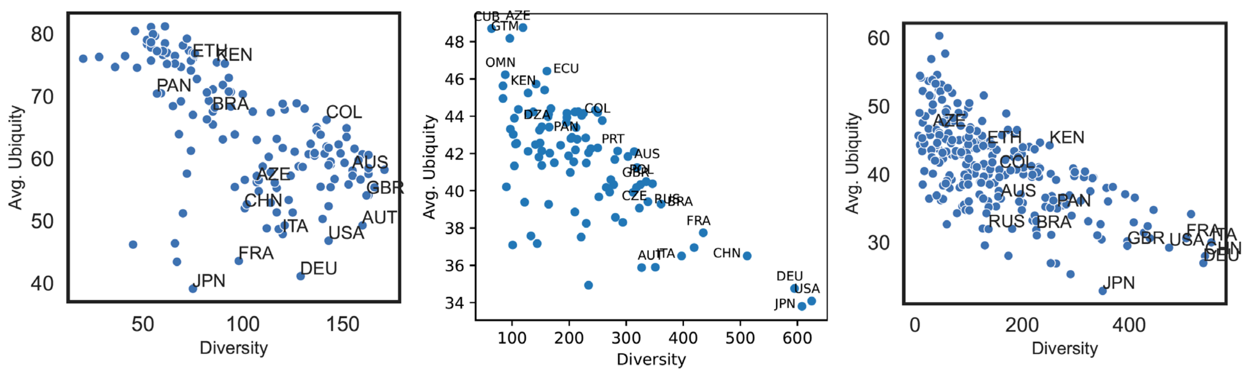
(a) Trade (b) Scientific Publications (c) Patents

Source: Own construction, using trade data from UN COMTRADE, scientific publications from OpenAlex, and patents compiled by WIPO. Number of unique categories in each matrix: scientific fields - 284, IPC classes - 650, HS product codes - 1248.

Countries with a high diversity of scientific publications, patents, and exports also tend to produce more unique outputs (Figure 6).[4] In each dimension, our diversity measure, which represents the number of outcomes in which a country has an RCA greater than one or holds a dominant position, is negatively correlated with the average ubiquity, which accounts for how widespread their revealed advantage outputs are.[5] This suggests that more diverse countries create products, scientific knowledge, and technologies that are less common globally.

Figure 7 shows the changes in average diversity over time from each of the three paradigms of publications, patenting, and trade. Global average diversity has increased by over 50% for scientific publications, by around 20% for patents, and has largely remained stable for trade over 2000-2020. From the accompanying paper (Hausmann et al., 2024), we know that the world has seen a much stronger convergence in publications compared to patenting or income. This development of countries with low participation in scientific publication activity could explain this steep rise in the diversity of publications, relative to patents and trade.

Figure 6: Diversity and Average Ubiquity, 2000 – 2020

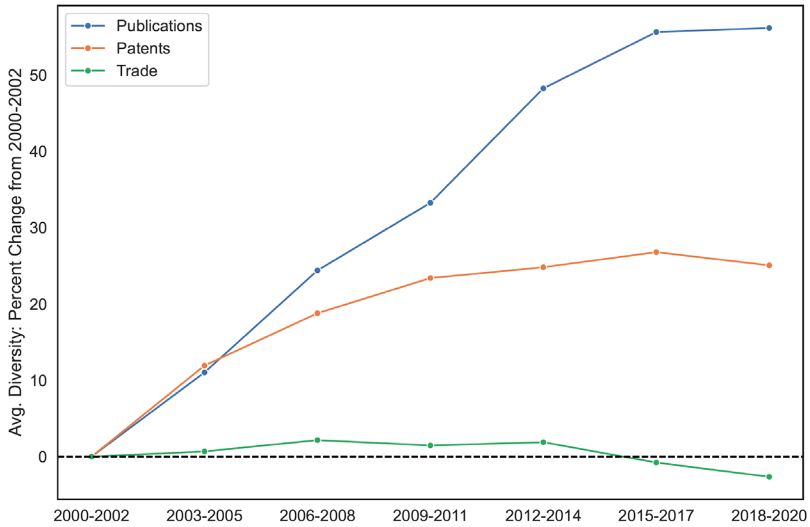


(a) Scientific Publications (b) Patents (c) Trade

Source: Own construction, using trade data from UN COMTRADE, scientific publications from OpenAlex, and patents compiled by WIPO.

[4] In the Appendix Section C, we also include Figure C.1, in which we only consider the most recent five-year period in our data (2015-2020). That figure gives a more recent and more static depiction of the relationship between diversity and average ubiquity.
[5] See the Appendix A to check the mathematical formulation and theoretical assumptions behind of our modified diversity and ubiquity measures for scientific publication and patent data.

Figure 7: Average Diversity

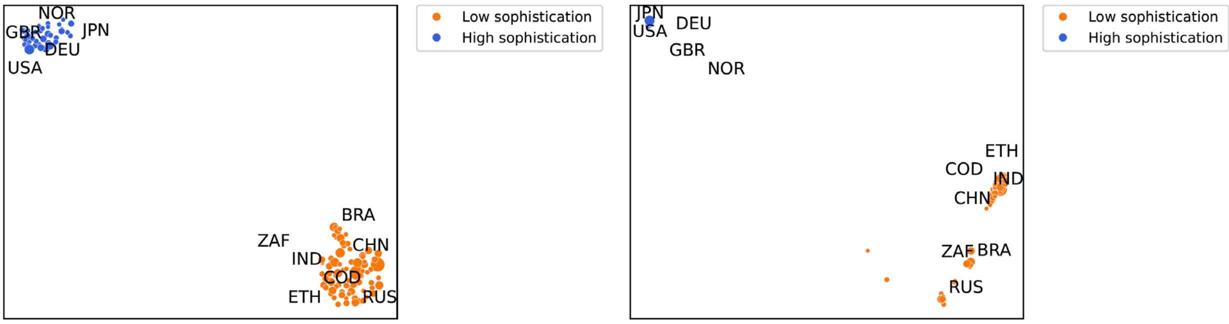


Source: Own construction, using trade data from UN COMTRADE, scientific publications from OpenAlex, and patents compiled by WIPO.

Next, we cluster the countries according to their production patterns in patents and scientific publications, aggregated for 2000-2020. When analyzing scientific and technological knowledge production, two distinct clusters of countries emerge.[6] Based on scientific publication patterns (Figure 8a), two remarkable groups can be identified: one group consists of advanced economies (e.g., Japan, Germany) with highly sophisticated publishing patterns, whereas the other group consists of low and middle-income countries (e.g., Brazil, Russia). This is similarly reflected in the country space derived from patenting patterns (Figure 8b).

[6] Please refer to the Appendix Section B.4 for more details on the clustering approach. Critically, clusters are not based on the volume of publications but on the types of scientific fields in which countries publish.

Figure 8: Country spaces, 2000-2020.

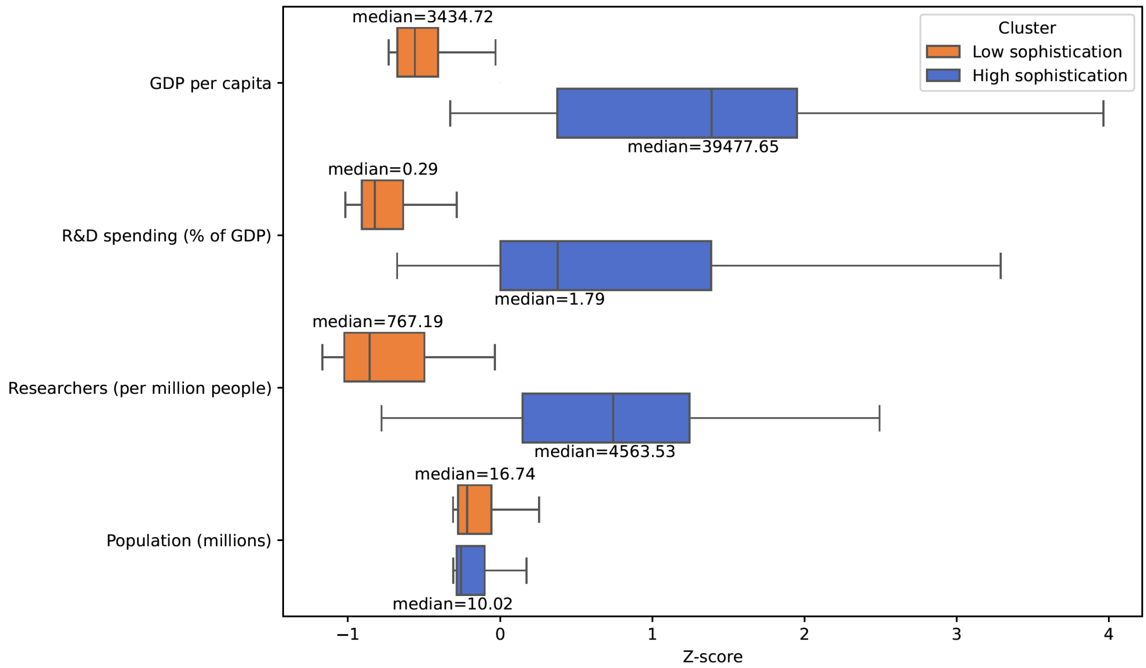


(a) Scientific Publications (b) Patents

Source: Own construction, using scientific publications from OpenAlex and patents compiled by WIPO.

Those advanced economies that engage mainly in complex activities also invest distinctively more in R&D and employ a significantly larger amount of researchers per million people (Figure 9), and, as expected, have a much higher GDP per capita. The differences in clusters do not seem to be associated with country populations. The scientific fields and technologies that are the best predictors of membership in one of the two country clusters (high or low sophistication) are listed in Tables C.1 and C.2 respectively (see appendix C). Although these lists appear logical, further research is required to ascertain whether transitioning into particular fields and technologies enhances sophistication (regarding ECI) and economic outcomes more effectively than others.

Figure 9: Country characteristics by Cluster



Source: Own construction, using data from the World Bank's World Development Indicators.

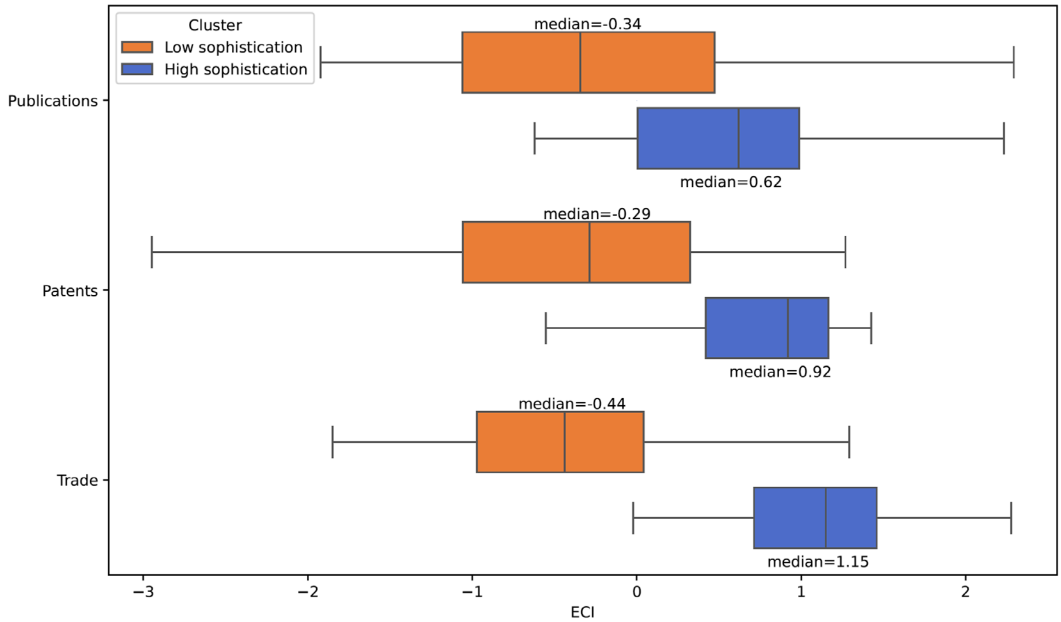
4 Complexity and Economic Growth

Next, we capture the extent of capabilities that are present in a country through the economic complexity index (Hidalgo & Hausmann, 2009; Hausmann et al., 2014). The basic idea behind this methodology is that the extent of capabilities is revealed through incorporating information about who makes what. The details of the complexity calculations are described in the mathematical appendix.[7]

First, let us examine the complexity patterns of the clusters discussed in the preceding section (Figure 8). In addition to differing in research inputs and economic outcomes (Figure 9), these clusters also differ significantly concerning their ECI, defined through publications, patents, or trade (Figure 10). Hence, countries with high complexity have different knowledge and innovation portfolios from countries with low complexity.

[7] Note that other studies have developed similar complexity metrics, under different names, for our considered innovation domains (Balland & Rigby, 2017; Stojkoski et al., 2023). For example, Balland & Rigby (2017) call "Knowledge Complexity Index" to the complexity metrics derived from patent data. In this paper, we denote all these metrics as Economic Complexity Indexes (ECIs), while mentioning their particular innovation dimension (scientific research, invention, and production).

Figure 10: Country ECI by Cluster



Source: Own construction, using trade data from UN COMTRADE, scientific publications from OpenAlex, and patents compiled by WIPO.

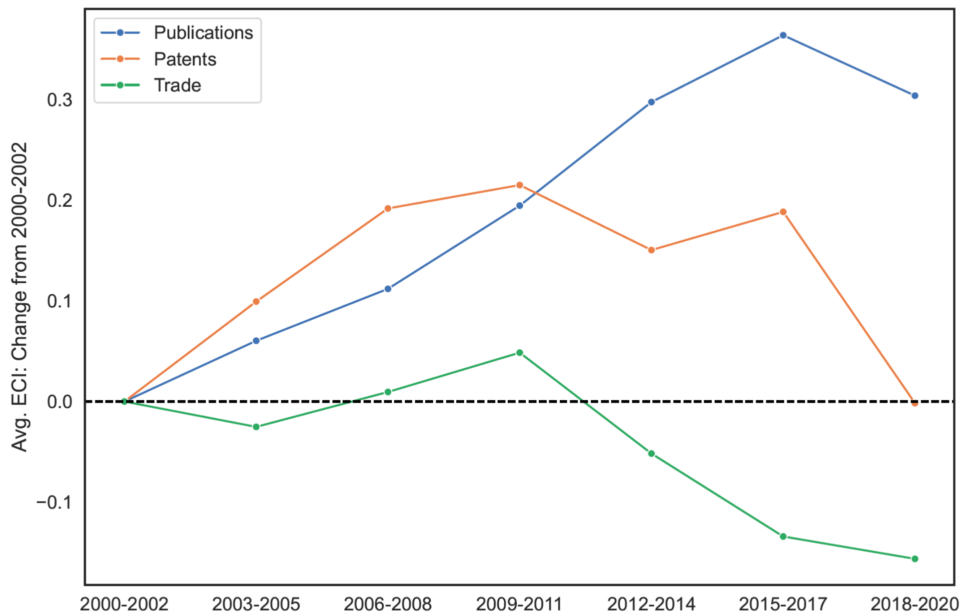
Trends in global economic complexity vary dramatically according to the innovation paradigm. Figure 11 shows the global average ECI over 2000-2020.[8] The average global ECI, as defined by scientific publications, has seen an increase of about 0.3 standard deviations over 2000-2020. The average ECI from patents rose by about 0.2 standard deviations from 2000-2010 and fell again after 2010, dramatically decreasing from 2015-2020. The average ECI for trade has largely remained stable, with a downward trend in recent years. These trends suggest that the average sophistication or complexity of scientific publications has increased over 2000-2020, whereas that of trade and patents has not. Figure 7 might explain this trend - it shows that the average diversity based on scientific publications has increased over time. This could explain the increasing average ECI based on publications relative to patents and trade.

We also examine country-specific changes in ECI for each paradigm to scrutinize the factors influencing these variations. Figure 12 shows the ECI in 2000 and 2020 for each

[8] To capture the dynamics in ECI, we stack presence-absence matrices, *M*, on top of each other and apply the complexity algorithm. Certainly, this would create a single complexity measure for products, but each country-year pair would have its own ECI values, which would allow us to study the ECIs' temporal dynamics.

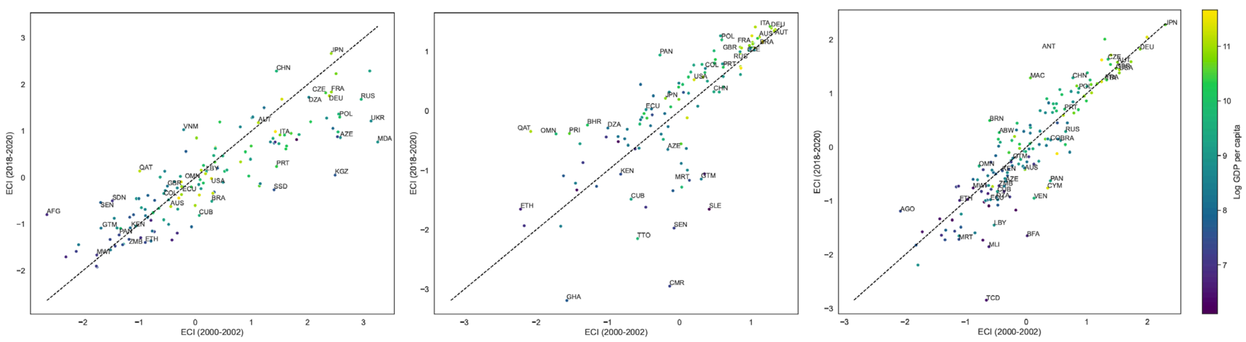
country. Income and complexity are correlated in each of the three paradigms. Most countries have seen an increase in publications and patenting complexity, but some countries have experienced big drops in patents-based economic complexity, which might explain the downturn in the global average complexity based on patents in recent years.

Figure 11: Average ECI



Source: Own construction, using trade data from UN COMTRADE, scientific publications from OpenAlex, and patents compiled by WIPO.

Figure 12: Change in ECI - 2000-2020

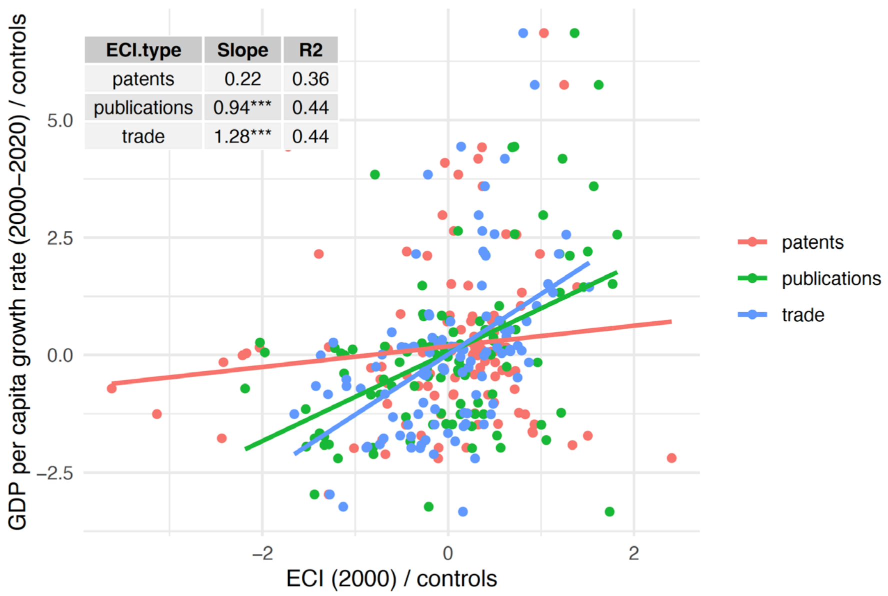


(a) Scientific Publications (b) Patents (c) Trade

Source: Own construction, using trade data from UN COMTRADE, scientific publications from OpenAlex, and patents compiled by WIPO.

Hausmann et al. (2014) and Hidalgo & Hausmann (2009) found that economic complexity is strongly positively related to economic growth. This also holds true for the ECI in publications, trade, and, although partially, for the ECI patents (Figure 13). We find that a standard deviation increase in ECI for the year 2000 is associated with an increase in the GDP per capita growth rate between 2000 and 2020 of around 94% for the scientific publications ECI; 128% for trade ECI; and 22% for patents ECI (although this coefficient is not significant). These results only control for the initial GDP per capita levels.

Figure 13: ECI vs. Growth



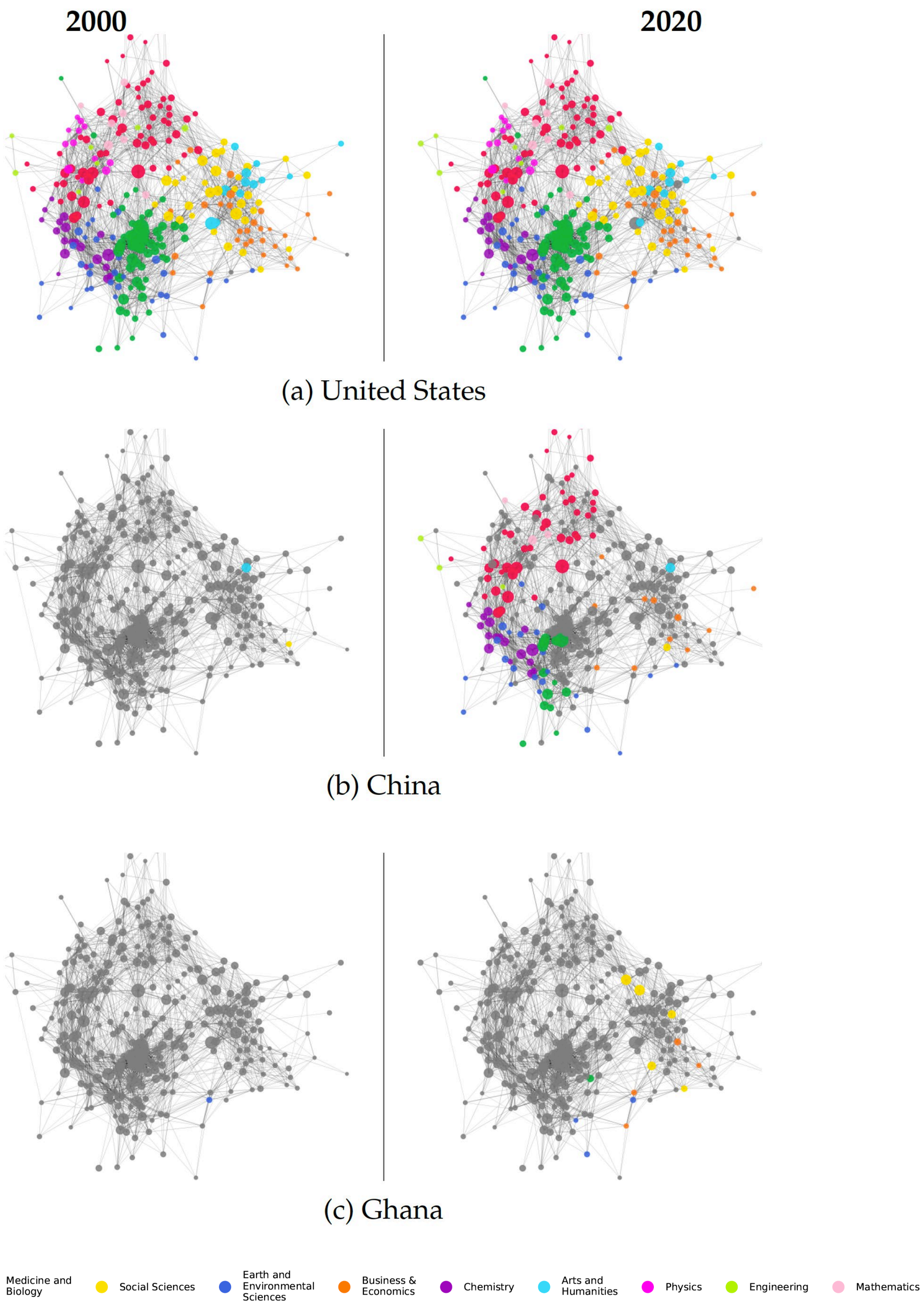
Source: Own construction, using trade data from UN COMTRADE, scientific publications from OpenAlex, and patents compiled by WIPO.

5 Path dependency

Accumulated productive know-how and capabilities shape the adjacent possible for countries. Structural transformation is a path-dependent process because economies tend to diversify incrementally, moving into related activities. Leap-frogging, hence, is unlikely as the required new capabilities would have to be acquired first. As a result of related diversification, economies tend to specialize in different areas.

To reveal the relatedness between scientific fields and technology classes, we build proximity matrices as in the product space (Hidalgo et al., 2007). We use measures of know-how overlaps such as co-occurrence in location, applicants, or co-citation patterns as described in Appendix Section A. The resulting examples of field space and technology spaces are shown in Figures 14 and 15, respectively. Interested readers can consult some examples of product spaces in Figure C.2, in Appendix Section C.

Figure 14: Evolution of the scientific field space for selected countries



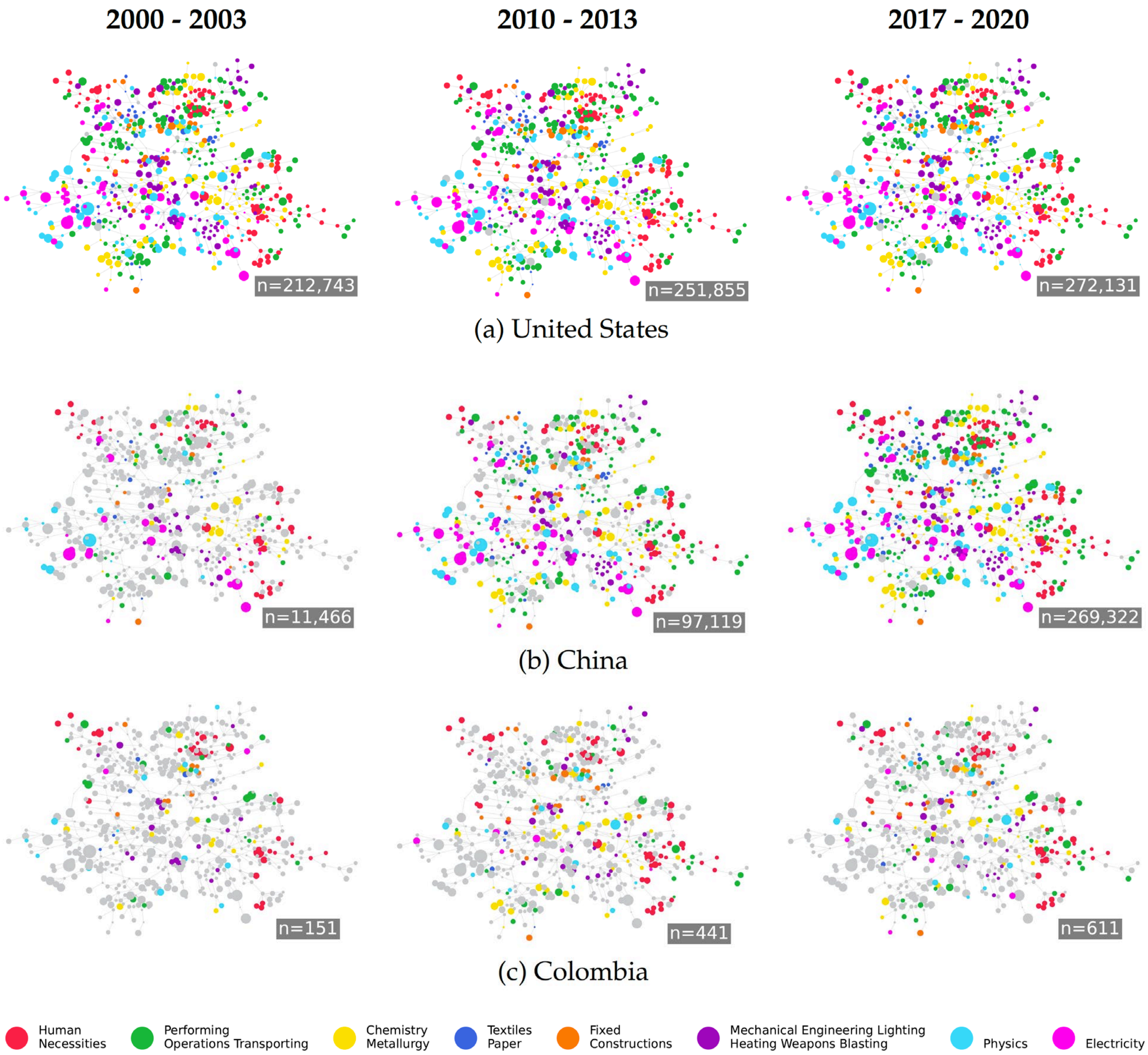
(a) United States

(b) China

(c) Ghana

Source: Own construction. Nodes, representing scientific fields, are sized by the number of publications and colored by major academic fields. Edges weights are based on the co-occurrence of fields in which authors publish (minimum conditional probability). Nodes displayed: 280, Edges displayed: 2707.

Figure 15: Evolution of the technology space for selected countries



Source: Own construction. Nodes, representing technologies at the IPC4 level, are sized by their share in the world's patents and colored by major technology groups. They are visualized in such a way that distinct clusters of related technologies can be identified. Edges weights are based on the co-occurrence of technologies in the same patent family. Nodes displayed: 648, Edges displayed: 736.

In Figure 14, each node corresponds to a scientific field, and colors encode the broad fields. The edges capture whether two fields share authors more than would be expected if those authors were shared randomly, directly capturing know-how overlap.[9] We can show

[9] We only show the network's maximum spanning tree and strong edges with proximity over 0.06.

each country on this network by coloring the nodes in which that country has a significant presence. As shown in Figure 14a, the United States publishes scientific papers in a wide variety of fields, ranging from Engineering and Computer Science to Business and Economics. This hardly changes over time - in 2020, its scientific portfolio looked very similar to that of its 2000 portfolio. Conversely, China is highly specialized in Engineering and Computer Science but much less present in other fields. To the extent that China's scientific production is included in our data, they rapidly specialized in this area over the past 15 years (Figure 14b).[10] Ghana is neither of the above. It publishes only in a few areas of Business & Economics.

These patterns are also reflected in patenting. Just like in the scientific field space, we build the technology space, where two technology classes are connected if those two technology classes often co-occur on the same patent. Looking at the technology space, the United States patents in a wide variety of fields - ranging from Mechanical Engineering to Human Necessities (Figure 15a). This hardly changes between 2000 and 2020. China's position in the same technology space was much sparser earlier on: in the 2000s, patents were concentrated in fewer areas - mainly in Mechanical Engineering and Human Necessities (Figure 15b). Over time, the diversity of fields increases and the technology space is much less sparse in 2020. Contrary to scientific publications, China has greatly diversified in patenting activity over time. Colombia patented only in a few areas in the 2000s, mainly focused on Human Necessities (Figure 15c). Over time, it slowly expanded its scope, mainly by diversifying into activities in Human Necessities it had not specialized in before, as well as gradually diversifying into Performing Operations and Transportation.

We formally test these patterns of related diversification in "density regressions," in which density measures the extent to which an activity is surrounded by related activities in a country. The expectation is that the more related activities there are for a particular activity, the more likely that activity is to appear because of similar capabilities. At the same time, unrelated

[10] Another potential explanation of China's specialization in Engineering and Computer science is related to OpenAlex's biases in coverage, mentioned in section 2 and the appendix B.2. OpenAlex may have low coverage of papers written in Chinese and Chinese journals in fields other than engineering and computer science, which may affect our measurement of China's specialization patterns.

activities are more likely to disappear. These patterns would align with similar density regressions of other studies (Hidalgo et al., 2007; Neffke et al., 2011; Balland et al., 2019; Balland & Boschma, 2022; Li & Neffke, 2023). The density variables are built as described in the mathematical formulations in the Appendix A. As controls, we either include country and field fixed effects or "radial growth" values as in Hausmann et al. (2022), which captures the overall growth of the location or the field in the considered time period. The dependent variable is the growth of RCA or the entry and exit of countries from technology classes or scientific fields. Entry events are defined as instances where the RCA surpasses 1 after starting from below an RCA level of 0.1. Conversely, exit events are defined as cases when an RCA transitions from above 1 to below 0.1.

Indeed, Table 1 shows that a country's existing scientific portfolio is strongly predictive of not just the entry of new scientific fields but also their exit and growth. In column 1, we regress the growth in RCA just with the current density level. The coefficient is highly significant and positive. In column 2, we add controls related to the base year only. In column 3, we use additional controls for the radial growth for the location and the field to capture the overall growth in those entities. In column 4, we add country and scientific field fixed effects, which capture all uncontrolled heterogeneities at the country and scientific field level. In all columns between 2 and 4, the coefficient of density variable is highly significant. In column 5, we run a logit regression to see whether the density variable is predictive of entry. In column 6, we do the same for the exits. As shown in columns 5 and 6 of Table 1, similar scientific activities are more likely to appear and grow when density is high and exit when density is low. Hence, existing activities are a strong predictor of future activities, with future activities likely being related to existing activities in which capabilities can be redeployed.

These results also hold for technological progress (Table 2). In column 1, we regress the current RCA on the density variable. The density variable is highly correlated with the current RCA. In columns 2 and 3, we use the controls explained above for Table 1 to predict the growth in RCA in technology classes. In column 4, we use country and technology fixed effects. Columns 2 to 4 validate that the density variable is predictive of future growth. As shown in columns 5 and 6 of Table 2, entry and growth are more likely when related technologies are present, while technologies with few related technologies around are more likely to disappear.

Table 1: Density regressions for scientific publications

	RCA	Growth			Appearance (Logit)	Disappearance (Logit)
	(1)	(2)	(3)	(4)	(5)	(6)
(Intercept)	−0.11***	−3.46***	−2.27***			
	(0.01)	(0.11)	(0.10)			
Density (IHS)	2.53***	0.59***	0.55***	1.16***	18.19***	−13.16***
	(0.02)	(0.06)	(0.06)	(0.20)	(2.12)	(2.10)
Base year value (IHS)		−0.14***	−0.08***	−0.30***		−1.93***
		(0.01)	(0.00)	(0.02)		(0.14)
Population (Log)		0.17***	0.10***			
		(0.00)	(0.00)			
Base year product total (Log)		0.15***	0.07***			
		(0.01)	(0.00)			
Location mean RCA		−0.03	−0.10***			
		(0.02)	(0.02)			
Radial product growth			6.38***			
			(0.12)			
Radial location growth			4.84***			
			(0.11)			
Num. Obs.	30 969	14 819	14 819	14 819	15 738	6326
Adj. R2	0.29	0.10	0.33	0.46	0.17	0.42
Country FE				X	X	X
Scientific Field FE				X	X	X

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Own estimations using scientific publications from OpenAlex and population and income data from the World Bank's World Development Indicators (WDI). Notes: The proximity matrix is computed from the cooccurrence of scientific fields in the same scientific publication. Logit estimates for appearance and disappearance models, and OLS estimates for the remaining analyses. We applied Inverse Hyperbolic Sine (IHS) and Logarithmic (Log) transformations as indicated next to the variable names. See the appendix B for more details on the data and methods.

Table 2: Density regressions for patents

	RCA	Growth			Appearance	Disappearance
	(1)	(2)	(3)	(4)	(5)	(6)
(Intercept)	0.04***	−3.05***	−2.34***			
	(0.00)	(0.09)	(0.09)			
Density (IHS)	2.09***	1.22***	1.69***	3.32***	9.55***	−15.11***
	(0.01)	(0.07)	(0.07)	(0.56)	(1.73)	(2.91)
Base Year Value (IHS)		−0.17***	−0.18***	−0.46***		−0.96***
		(0.01)	(0.01)	(0.03)		(0.10)
Population (Log)		0.16***	0.11***			
		(0.00)	(0.00)			
Base Year Product Total (Log)		0.15***	0.13***			
		(0.01)	(0.01)			
Country Mean RCA		−0.42***	−0.43***			
		(0.02)	(0.02)			
Radial Product Growth			6.32***			
			(0.14)			
Radial Location Growth			5.91***			
			(0.16)			
Num. Obs.	132 924	23 315	23 129	23 315	89 739	12 034
R2	0.14	0.08	0.20	0.34	0.27	0.52
Country FE				X	X	X
IPC4 Technology Class FE				X	X	X

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Own estimations using patents compiled by WIPO and population and income data from the World Bank'sWorld Development Indicators (WDI). Notes: The proximity matrix is computed from the co-occurrence of technology classes in the patents filed by an applicant. Logit estimates for appearance and disappearance models, and OLS estimates for the remaining analyses. We applied Inverse Hyperbolic Sine (IHS) and Logarithmic (Log) transformations as indicated next to the variable names. See the appendix B for more details on the data and methods.

6 Can the rest of the world achieve complex technologies?

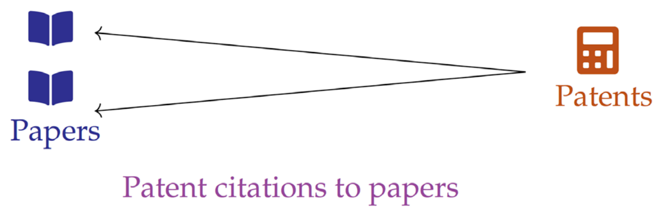
Countries are thus specialized in very different areas when it comes to trade, patents, and scientific publications. However, do these areas relate to each other? How do scientific capabilities, for instance, translate into economic capabilities? Moreover, how does scientific output relate to patenting? They may not be directly correlated and may not co-evolve together. Furman et al. (2002) show that patenting activity across countries correlates with scientific publications, but not every publication necessarily leads to patenting. At the regional level, Balland & Boschma (2022) find that scientific capabilities in domains predict the development of related new technologies in the corresponding domain in regions.

In this section, we study the association between technologies and academic fields and between technologies and products using cross-domain linkage tables (Figure 16). We rely on information from patent citations to scientific papers, obtained from Marx & Fuegi (2020), to link technologies and academic fields. Moreover, we combine information on Colombian imports (DIAN, 2023) with worldwide patents, by matching data at the firm level, to relate technologies and products. More details on the methods to build these cross-domain linkage tables are included in the Appendix Sections B.5 and B.6.

Figure C.16: Cross-domain linkage tables

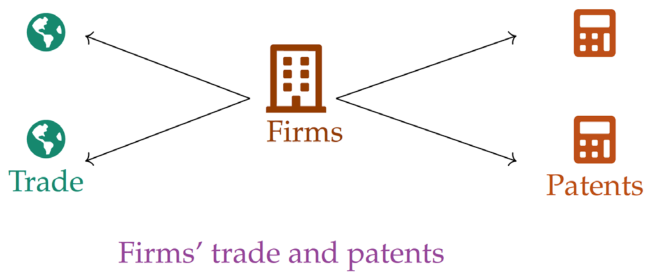
Fields	IPC4
F1	T1
⋮	⋮

=



HS Code	IPC4
H1	T1
⋮	⋮

=



Source: Own construction.

Table 3 examines the association between the number of scientific publications in various fields and the number of patent families in future years in related technology classes. The extent to which technology classes are related to scientific publications is estimated by analyzing citations of patent families in various IPC technology classes to scientific fields in OpenAlex (similar in approach to, e.g., Shin et al., 2023). The regressions show that countries are more likely to diversify in technologies related to their existing scientific capabilities. Although this result does not allow the establishment of a causal relationship, it is robust across models and holds when including country and technology fixed effects.

Table 3: Publications and Related Patenting

	Patents at t-5 (IHS)	Growth in Patents (t-5 to t, IHS)			
	1	2	3	4	5
(Intercept)	−0.01	0.12***	0.12***		
	(0.02)	(0.02)	(0.02)		
Patents Estd. t-5 (IHS)	0.51***	0.15***	0.35***	0.29***	0.20***
	(0.00)	(0.00)	(0.00)	(0.02)	(0.03)
Patents at t-5 (IHS)			−0.40***	−0.45***	−0.64***
			(0.01)	(0.03)	(0.04)
Num. Obs.	23 923	23 923	23 923	23 923	23 923
Adj. R2	0.49	0.06	0.18	0.46	0.62
Country FE				X	X
Technology Class FE					X

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Own estimations, using scientific publications from OpenAlex, patents compiled by WIPO, and patent citations to papers from Marx & Fuegi (2020). We applied Inverse Hyperbolic Sine (IHS) transformations as indicated next to the variable names. See Appendix B for more details on the data methods.

Similarly, we can investigate the link between countries' trade portfolios and the probability of entering new technologies. To accomplish this, we estimate the relatedness between technologies and products by linking worldwide exporter data to Colombia (Colombian imports) to worldwide patent data. Those that patent in semiconductors, for instance, also tend to export products related to semiconductors more often (to Colombia). Hence, technology codes on patents can be related to product codes of products by incorporating the extent to which products are disproportionally related to certain technologies. Using this methodology, Table 4 shows that countries are more likely to enter new technologies related to products they are currently trading. Again, without claiming causality, these results are robust to the inclusion of country and country-technology fixed effects.

Table 4: Trade and Related Patenting

	Patents at t-5 (IHS)	Growth in Patents (t-5 to t, IHS)			
	1	2	3	4	5
(Intercept)	-4.75^{***}	0.78^{***}	-0.78^{***}		
	(0.08)	(0.08)	(0.08)		
Patents Estd. t-5 (IHS)	0.35^{***}	0.01^{*}	0.13^{***}	0.06^{***}	0.09^{***}
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Patents at t-5 (IHS)			-0.33^{***}	-0.25^{***}	-0.68^{***}
			(0.01)	(0.02)	(0.04)
Num. Obs.	19 856	19 856	19 856	19 856	19 856
Adj. R2	0.26	0.00	0.13	0.39	0.61
Country FE				X	X
Technology Class FE					X

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Own estimations, using trade data from UN COMTRADE, patents compiled by WIPO and Google, and detailed import transactions from DIAN (2023). We applied Inverse Hyperbolic Sine (IHS) transformations as indicated next to the variable names. See Appendix B for more details on the data methods.

7 Conclusions and policy implications

This paper presents evidence that the development of capabilities of countries is correlated across economic activities, science, and technologies. Existing activities in each area are predictive of future activity in other areas. Our results align with similar findings of technological diversification of countries by Petralia et al. (2017), as well as the relationship at the regional level between science and technology by Balland & Boschma (2022) and Shin et al. (2023). They also reinforce findings of case studies analyzing the co-evolution of industrial and academic domains in regions by Lehmann & Menter (2016) and Kenney & Mowery (2020). This paper also incorporates economic activities - measured through trade patterns - into this framework.

In this paper, we mainly highlighted the related diversification opportunities for countries. However, Coniglio et al. (2021) and Pinheiro et al. (2022) also highlight the importance of unrelated diversification for faster growth. Especially for countries with a limited number of capabilities, a sustained unrelated diversification to a new innovation field could trigger further diversification opportunities. Nevertheless, the related diversification measure's highly predictive power of disappearances in Tables 1 and Tables 2 show that many unrelated diversification events often become futile. Hence, there is a greater risk of failure associated with unrelated diversification events, and this risk must be carefully assessed by entrepreneurs and policymakers.

Policymakers must consider the co-evolution of different cognitive domains and their interdependencies. Promoting, for instance, the acceleration of certain scientific domains through public funding may induce positive externalities that increase the likelihood of new technological capabilities emerging related to those scientific domains. Similarly, the emergence of a technological capability is correlated with countries' past exports in related domains. Hence, developing complex technologies can leverage existing scientific and industrial capabilities. Disentangling how these capabilities interact in a given location and identifying their most binding constraints to grow may boost the presence of positive externalities and the development of a resilient innovation ecosystem.

As highlighted in the accompanying paper (Hausmann et al., 2024), fostering knowledge infusion from elsewhere is essential to foster diversification. Neffke et al. (2018) found that regional structural change comes not so much from incumbent firms but from entrepreneurs and expanding existing firms, particularly when they come from elsewhere. Similarly, more recently, Miguelez & Morrison (2023) find that immigrant inventors foster technological diversification by developing new technological specializations, which, in turn, are transferred from the home country to the host region. Increasing evidence highlights the importance of the movement of brains - for instance, through migration (Morrison, 2023) - for the structural transformation of economies.

References

Ansorge, L. (2023). Hidden limitations of analyses via alternative bibliometric services. *Scientometrics*, *128*(3), 2031–2033. doi: 10.1007/s11192-022-04626-w

Arrow, K. (1962). Economic welfare and the allocation of resources for invention. In *The rate and direction of inventive activity: Economic and social factors* (p. 609-626). Princeton University Press.

Balassa, B. (1965). Trade liberalisation and “revealed" comparative advantage. *The Manchester School*, *33*(2), 99–123.

Balland, P.-A., & Boschma, R. (2022). Do scientific capabilities in specific domains matter for technological diversification in European regions? *Research Policy*, *51*(10), 104594. doi: 10.1016/j.respol.2022.104594

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

Balland, P.-A., & Rigby, D. (2017). The geography of complex knowledge. *Economic Geography*, *93*(1), 1–23.

Breschi, S., & Lissoni, F. (2001). Knowledge spillovers and local innovation systems: A critical survey. *Industrial and Corporate Change*, *10*(4), 975–1005.

Burbidge, J. B., Magee, L., & Robb, A. L. (1988). Alternative Transformations to Handle Extreme Values of the Dependent Variable. *Journal of the American Statistical Association*, *83*(401), 123–127. doi: 10.2307/2288929

Bustos, S., & Yildirim, M. A. (2023). *Uncovering trade flows.* (Unpublished mimeo, available upon request.)

Chen, J., & Roth, J. (2023). *Logs with zeros? Some problems and solutions* (No. arXiv:2212.06080). arXiv. doi: 10.48550/arXiv.2212.06080

Cohen, W. M., Nelson, R., & Walsh, J. P. (2000). *Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)* (Working Paper No. 7552). Cambridge, MA, USA: National Bureau of Economic Research.

Coniglio, N. D., Vurchio, D., Cantore, N., & Clara, M. (2021). On the evolution of comparative advantage: path-dependent versus path-defying changes. *Journal of International*

Economics, *133*, 103522.

Cowan, R., David, P., & Foray, D. (2000). The explicit economics of knowledge codification and tacitness. *Industrial and corporate change*, *9*(2), 211–253.

DIAN. (2023). *Bases estadísticas de comercio exterior - imports data.* Retrieved from https://www.dian.gov.co/dian/cifras/Paginas/Bases-Estadisticas-de-Comercio-Exterior-Importaciones-y-Exportaciones.aspx (Accessed on July 1st, 2023)

Furman, J. L., Porter, M. E., & Stern, S. (2002). The determinants of national innovative capacity. *Research policy*, *31*(6), 899–933.

Golden, J. M. (2007). "Patent trolls" and patent remedies. *Texas law review*, *85*(7), 2111.

Gruber, J., & Johnson, S. (2019). *Jump-starting america: How breakthrough science can revive economic growth and the american dream*. Hachette UK.

Harabi, N. (1995). Appropriability of technical innovations an empirical analysis. *Research Policy*, *24*(6), 981–992.

Harvard's Growth Lab. (2023). *The Atlas of Economic Complexity.* Retrieved 2023-09-26, from https://atlas.cid.harvard.edu/

Harzing, A.-W., & Alakangas, S. (2017). Microsoft Academic: Is the Phoenix Getting Wings? *Scientometrics*, *110*(1), 371–383. doi: 10.1007/s11192-016-2185-x

Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., Simoes, A., & Yıldırım, M. A. (2014). *The atlas of economic complexity: Mapping paths to prosperity*. The MIT Press.

Hausmann, R., Stock, D. P., & Yıldırım, M. A. (2022). Implied comparative advantage. *Research Policy*, *51*(8), 104143.

Hausmann, R., Yildirim, M. A., Chacua, C., Hartog, M., & Matha, S. G. (2024). Innovation policies under economic complexity. *WIPO Economic Research Working Paper*(TO BE CONFIRMED).

Henkel, J. (2022). Licensing standard-essential patents in the IoT – A value chain perspective on the markets for technology. *Research policy*, *51*(10), 104600.

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

Hidalgo, C. A., Klinger, B., Barabási, A.-L., & Hausmann, R. (2007). The product space conditions the development of nations. *Science*, *317*(5837), 482–487.

Huang, C.-K. K., Neylon, C., Brookes-Kenworthy, C., Hosking, R., Montgomery, L., Wilson, K., & Ozaygen, A. (2020). Comparison of bibliographic data sources: Implications for the robustness of university rankings. *Quantitative Science Studies*, *1*(2), 445–478. doi: 10.1162/ qss_a_00031

Jiao, C., Li, K., & Fang, Z. (2023). *How are exclusively data journals indexed in major scholarly databases? An examination of the Web of Science, Scopus, Dimensions, and OpenAlex* (No. arXiv:2307.09704). arXiv. doi: 10.48550/arXiv.2307.09704

Kenney, M., & Mowery, D. C. (2020). *Public universities and regional growth: Insights from the University of California*. Stanford University Press.

Lehmann, E. E., & Menter, M. (2016). University–industry collaboration and regional wealth. *The Journal of Technology Transfer*, *41*, 1284–1307.

Lemley, M. A., & Shapiro, C. (2007). Patent holdup and royalty stacking. *Texas law review*, *85*(7), 1991.

Levin, R. C., Klevorick, A. K., Nelson, R. R., Winter, S. G., Gilbert, R., & Griliches, Z. (1987). Appropriating the returns from industrial research and development. *Brookings papers on economic activity*, *1987*(3), 783–831.

Li, Y., & Neffke, F. (2023). *Evaluating the principle of relatedness: Estimation, drivers and implications for policy* (CID Research Fellows and Graduate Student Working Paper Series No. 146). Cambridge, MA: Center for International Development at Harvard University.

MacKinnon, J. G., & Magee, L. (1990). Transforming the Dependent Variable in Regression Models. *International Economic Review*, *31*(2), 315–339. doi: 10.2307/2526842

Martín-Martín, A., Thelwall, M., Orduna-Malea, E., & Delgado López-Cózar, E. (2021). Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: A multidisciplinary comparison of coverage via citations. *Scientometrics*, *126*(1), 871– 906. doi: 10.1007/s11192-020-03690-4

Marx, M., & Fuegi, A. (2020). Reliance on science: Worldwide front-page patent citations to scientific articles. *Strategic Management Journal*, *41*(9), 1572–1594.

Mazzucato, M. (2015). *The entrepreneurial state: debunking public vs. private sector myths* (Revised edition. ed.). New York, NY: PublicAffairs.

McInnes, L., Healy, J., Saul, N., & Großberger, L. (2018). Umap: Uniform manifold approximation and projection. *Journal of Open Source Software*, *3*(29), 861. doi:

10.21105/joss.00861

Mezzanotti, F., & Simcoe, T. (2023). *Innovation and appropriability: Revisiting the role of intellectual property* (Working Paper No. 31428). Cambridge, MA, USA: National Bureau of Economic Research.

Miguelez, E., & Morrison, A. (2023). Migrant inventors as agents of technological change. *The Journal of Technology Transfer*, *48*(2), 669–692.

Miguelez, E., Raffo, J., Chacua, C., Coda-Zabetta, M., Yin, D., Lissoni, F., & Tarasconi, G. (2019). Tied in: The global network of local innovation. *WIPO Economic Research Working Paper*, *58*.

Morrison, A. (2023). Towards an evolutionary economic geography research agenda to study migration and innovation. *Cambridge Journal of Regions, Economy and Society*, rsad013.

Mullahy, J., & Norton, E. C. (2022). *Why Transform Y? A Critical Assessment of Dependent-Variable Transformations in Regression Models for Skewed and Sometimes-Zero Outcomes* (Working Paper No. 30735). National Bureau of Economic Research. doi: 10.3386/w30735

Neffke, F., Hartog, M., Boschma, R., & Henning, M. (2018). Agents of structural change: The role of firms and entrepreneurs in regional diversification. *Economic Geography*, *94*(1), 23–48.

Neffke, F., Henning, M., & Boschma, R. (2011). How do regions diversify over time? industry relatedness and the development of new growth paths in regions. *Economic geography*, *87*(3), 237–265.

Nelson, R. R. (1959). The simple economics of basic scientific research. *The Journal of political economy*, *67*(3), 297–306.

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

Pinheiro, F. L., Hartmann, D., Boschma, R., & Hidalgo, C. A. (2022). The time and frequency of unrelated diversification. *Research Policy*, *51*(8), 104323.

Priem, J., Piwowar, H., & Orr, R. (2022). OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. *arXiv preprint arXiv:2205.01833*.

Scheidsteger, T., & Haunschild, R. (2022). Comparison of metadata with relevance for bibliometrics between Microsoft Academic Graph and OpenAlex until 2020. In *26th*

International Conference on Science and Technology Indicators. Granada, Spain. doi: 10.5281/ zenodo.6975102

Scheidsteger, T., Haunschild, R., & Bornmann, L. (2023). How similar are field-normalized scores from different free or commercial databases calculated for large German universities? In *27th International Conference on Science, Technology and Innovation Indicators (STI 2023).* International Conference on Science, Technology and Innovation Indicators. doi: 10.55835/ 6441118c643beb0d90fc543f

Shin, H., Kogler, D. F., & Kim, K. (2023). The relevance of scientific knowledge externalities for technological change and resulting inventions across European metropolitan areas. *Review of Regional Research*, 1–17.

Stiglitz, J. E. (1999). Knowledge as a global public good. In *Global public goods.* New York: Oxford University Press.

Stojkoski, V., Koch, P., & Hidalgo, C. A. (2023). Multidimensional economic complexity and inclusive green growth. *Communications Earth & Environment*, *4*(1), 130.

Thelwall, M. (2017). Microsoft Academic: A multidisciplinary comparison of citation counts with Scopus and Mendeley for 29 journals. *Journal of Informetrics*, *11*(4), 1201–1212. doi: 10.1016/j.joi.2017.10.006

Wang, K., Shen, Z., Huang, C., Wu, C.-H., Dong, Y., & Kanakia, A. (2020). Microsoft Academic Graph: When experts are not enough. *Quantitative Science Studies*, *1*(1), 396–413. doi: 10.1162/qss_a_00021

Wang, K., Shen, Z., Huang, C., Wu, C.-H., Eide, D., Dong, Y., . . . Rogahn, R. (2019). A Review of Microsoft Academic Services for Science of Science Studies. *Frontiers in Big Data*, *2*.

A Mathematical Formulations

All our observations are at the country level, denoted by c. We start with three types of country-level data, namely trade, innovation, and scientific outputs. For the trade data, we denote the value of trade by country c in industry (product) i with X_{ci}. For the innovation data, P_{ct} represents the number of patents in technology class t by country c. Finally, for the scientific output, S_{cf} is the number of publications by country c in scientific field f. Whenever needed, to illustrate the time dimension, we will add y as an index to show the year.

Below, we will use the international trade data to describe the relevant variables, and these definitions can easily be extended to other spaces. For our derived variables, such as the Economic complexity, we will use superscripts of X, P, or S to reflect the source of the data.

A.1 Who makes what

A.1.1 Comparative Advantage

Originally, many complexity variables relied on the binary presence/absence of industries in locations. These binary variables are generated via comparing the country's production level to a benchmark, which gives comparative advantage measures. The comparative advantage measures have the following common structure:

$$R_{ci} = X_{ci}/\hat{X}_{ci},$$

where $\hat{X}_{ci}$ is the expected level of trade value in the country. One of the most used comparative advantage measures, namely Balassa's Revealed Comparative Advantage (RCA) (Balassa, 1965), assumes that the expected level of trade of country c in industry i should be proportional to the share of country c in world exports. We denote the total exports of country c as X_c, total exports of product i as X_i and total trade in the world with X_W, Mathematically, these can be written as:

$$X_c \equiv \sum_i X_{ci}, \qquad X_i \equiv \sum_c X_{ci} \quad \text{and} \quad X_W \equiv \sum_c \sum_i X_{ci}.$$

With these in hand, we can write the expected value of the exports of country c in industry

i as:

$$\hat{X}_{ci} = X_i \frac{X_c}{X_W}.$$

Therefore, the RCA measure is defined as:

$$\text{RCA}_{ci} \equiv \frac{X_{ci}/X_i}{X_c/X_W} = \frac{X_{ci}/X_c}{X_i/X_W} = \frac{X_{ci}X_W}{X_cX_i}.$$

RCA measure has some peculiarities. For a given country c and for an industry i, the RCA measure always satisfies:

$$\text{RCA}_{ci} \leq \frac{X_W}{X_c} \quad \text{and} \quad \text{RCA}_{ci} \leq \frac{X_W}{X_i}.$$

For a large country like China, which accounts for $\frac{X_c}{X_W} = 13.3\%$ of the world trade, RCA values would be always smaller than approximately 7.5. But for a small country like Nepal, this share is $.000071$, putting the theoretical upper limit close to 14,000. To even out the field for all countries and products, we cap the RCA measure with the minimum of these theoretical limits. Mathematically, we define the RCA cap as:

$$\overline{\text{RCA}} \equiv \min\left\{\left\{\frac{X_W}{X_c}\right\}_{\forall c}, \left\{\frac{X_W}{X_i}\right\}_{\forall i}\right\}.$$

And we replace RCA values above $\overline{\text{RCA}}$ with the cap value.

Another comparative advantage value uses the population share of a country c while calculating the expected trade. Mathematically, this accounts to:

$$\hat{X}_{ci} = X_i \frac{\text{pop}_c}{\text{pop}_W},$$

where pop_c captures the population of country c and pop_W captures that of the world. The Revealed per capita advantage (RpCA) is defined as:

$$\text{RpCA}_{ci} \equiv \frac{X_{ci}/X_i}{\text{pop}_c/\text{pop}_W}.$$

Similar capping to the RpCA can be applied using the population shares of countries and trade shares of the products.

A.1.2 Binary presence/absence matrix

Many complexity calculations require defining a presence/absence matrix. Comparative advantage measures, such as RCA, provide a good basis for such calculations. In particular, we assume that the country has all the necessary capabilities (letters) for making a product if it competitively exports the product with a comparative advantage value larger than 1. Mathematically, we define the country presence/absence matrix, M, as follows:

$$M_{c,i} = \begin{cases} 1 & \text{if} \quad \mathrm{RCA}_{ci} \geq 1 \\ 0 & \text{otherwise.} \end{cases}$$

This simple definition has a caveat, especially for innovation and scientific output datasets. As shown in Figure C.4b, the inequality in patenting and scientific publications is high. For example, a country like the United States is active in almost all technology classes or scientific fields, but because of the nature of the RCA variable, some of them would be considered as absent because of the threshold of 1 even though the US could be among the top innovators or publishers in the field. To circumvent this issue, we add an additional criterion allowing us to assign a presence value if the country is among the top-ranked countries in a technology class or a scientific field. The effective number of countries active in a field could be defined as the inverse Herfindahl–Hirschman Index (HHI):

$$\frac{1}{n_t} \equiv HHI_t \equiv \sum_c \left(\frac{P_{ct}}{P_t}\right)^2$$

where P_t is the total number of patents in the technology class globally. Let's define rank_{ct} as the rank of country c in technology class t. Then, the updated M^P matrix is defined as:

$$M^P_{c,t} = \begin{cases} 1 & \text{if} \quad \mathrm{RCA}_{ct} \geq 1 \text{ or } \mathrm{rank}_{ct} \leq n_t \\ 0 & \text{otherwise.} \end{cases}$$

A similar modification is done for the presence/absence matrix for scientific output, M^S.

A.2 Economic Complexity and Product Complexity Indexes

With binary presence/absence matrix -M- in hand, we can define the economic complexity metrics. First, we define diversity as the number of products that a country has a presence in. Mathematically, we define diversity, $k_{c,0}$, as:

$$k_{c,0} \equiv \sum_i M_{ci}.$$

Similarly, the ubiquity of a product is defined as the number of countries making the product:

$$k_{i,0} \equiv \sum_c M_{ci}.$$

Highly complex countries have relatively high diversity, and highly complex products are made by fewer countries exhibiting lower ubiquity levels. These measures could be considered the zeroth order approximation to the complexity levels (hence the zero in their subscript).

The Economic Complexity Index (ECI) and the Product Complexity Index (PCI) are calculated as refinements to diversity and ubiquity measures. Hidalgo & Hausmann (2009) introduce these measures as an iterative process with the following:

$$k_{c,n} = \frac{1}{k_{c,0}} \sum_i M_{ci}\, k_{i,n-1} \quad \text{and} \quad k_{i,n} = \frac{1}{k_{c,0}} \sum_c M_{ci}\, k_{c,n-1}.$$

This iterative process has a trivial solution where all complexity variables are equal to each other. If we focus on the component driving the difference between countries, this process results in the following specification that gives us ECI and PCI:

$$\mathrm{ECI}_c = \frac{\gamma}{k_{c,0}} \sum_i M_{ci}\, \mathrm{PCI}_i \quad \text{and} \quad \mathrm{PCI}_i = \frac{\gamma}{k_{c,0}} \sum_c M_{ci}\, \mathrm{ECI}_c,$$

where γ is a constant that will be determined below. We can write these equations in terms of matrix equations. First, we define the matrix D (U) as the diagonal matrix whose diagonal elements correspond to the diversity (ubiquity). Hence:

$$\mathrm{ECI} = \gamma\, D^{-1}\, M\, \mathrm{PCI} \quad \text{and} \quad \mathrm{PCI} = \gamma\, U^{-1}\, M^{\dagger}\, \mathrm{ECI}.$$

Combining these two equations yields:

$$\mathrm{ECI} = \gamma^2\, D^{-1}\, M\, U^{-1}\, M^{\dagger}\, \mathrm{ECI}.$$

Mathematically, we define ECI as the eigenvector corresponding to the second largest eigenvector (λ_2) of the matrix $\widetilde{M} = D^{-1}\, M\, U^{-1}\, M^{\dagger}$ with $\gamma = (\lambda_2)^{-1/2}$.

A.3 Inferring Capability Overlaps

A.3.1 Relatedness

Relatedness or proximity measure captures the capability overlap between products, between technology classes, or between scientific fields. In the international trade data, we infer the capability overlaps through the co-location of exports. In particular, we measure the probability of a country exporting a product i given that the country already exports i'. To minimize the error, we take the minimum of these conditional probabilities between these products. Mathematically, the proximity between to products, i and i' is:

$$\phi_{ii'} = \frac{\sum_c M_{ci} M_{ci'}}{\max(k_{i,0}, k_{i',0})}.$$

In the innovation and scientific output data, we can also use more direct measures of knowledge overlaps, in addition to co-location based measures. In the first approach, we use the co-occurrence of applicants in patents or co-authorship of scientific articles as an indicator of relatedness. Suppose $A_{tt'}$ is the number of applicants with patents in both technology classes t and t'. Then, we can write the proximity as:

$$\Phi_{tt'} = \frac{A_{t,t'}}{\sum_{t'} A_{t,t'}}.$$

As a second alternative, we can use citations between patents or scientific publications to quantify the overlap between productive knowledge required in different technology classes or scientific fields. For example, let's denote the total citations from technology class t' to technology class t with $C_{t,t'}$. Then, we can define a relatedness metric based on co-citations as:

$$\Phi_{tt'} = \frac{C_{t,t'}}{\sum_{t'} C_{t,t'}}.$$

Our third alternative is to use the fact that patents can be classified into multiple technology classes or scientific articles could be classified into multiple scientific fields. Suppose the number of patents classified into technology classes t and t' is $N_{tt'}$. Then the relatedness can be defined as:

$$\Phi_{tt'} = \frac{N_{t,t'}}{\sum_{t'} N_{t,t'}}.$$

A.3.2 Density

We infer the overlap between capabilities present in a location and capabilities required for a product, a technology class, or a scientific field through a measure called density. The density measure is calculated as the share of "relatedness" present in a location around a product. Mathematically, we can write the density as defined in Hidalgo et al. (2007):

$$D_{ci} = \frac{\sum_{i'} M_{ci'} \phi_{ii'}}{\sum_{i'} \phi_{ii'}}.$$

The numerator captures the portion of the related products that are present in the country. If the country makes products in industries with high relatedness, we expect the density value to be large. One caveat of the density measure is that if the country makes many products, the density also increases.

We can also write a density for the continuous RCA-like measures, which Hausmann et al. (2022) define as implied comparative advantage. This measure is calculated as the weighted average of RCAs of a country in related products. Mathematically, the implied comparative advantage measure can be written as:

$$\hat{R}_{ci} = \frac{\sum_{i'} \text{RCA}_{ci'} \phi_{ii'}}{\sum_{i'} \phi_{ii'}}.$$

A.3.3 Metrics for Density Regressions

To measure the growth of comparative advantage of a country c in product i, we use the following growth measure:

$$\text{Growth}_{ci,y-y'} = \frac{r_{ciy'}}{r_{ciy}} - 1$$

Where r_{ciy} is the RCA of country c in product, scientific field, or technology class i at time y.

We define the appearance of a product i in country c as:

$$\text{Appearance}_{ci,y-y\prime} = \begin{cases} 1 & \text{if} \quad r_{ciy} \leq 0.1 \text{ and } r_{ciy\prime} \geq 1 \\ 0 & \text{if} \quad r_{ciy} \leq 0.1 \text{ and } r_{ciy\prime} < 1 \\ \text{undefined} & \text{otherwise} \end{cases}$$

and disappearance as:

$$\text{Disappearance}_{ci,y-y\prime} = \begin{cases} 1 & \text{if} \quad r_{ciy} \geq 1 \text{ and } r_{ciy\prime} \leq 0.1 \\ 0 & \text{if} \quad r_{ciy} \geq 1 \text{ and } r_{ciy\prime} > 0.1 \\ \text{undefined} & \text{otherwise} \end{cases}$$

We use the Inverse Hyperbolic Sine (IHS) transformation instead of the log transformation. Density can be zero, and growth can often be zero or negative. Unlike the log transformation, the IHS transformation is defined also for values that are not strictly positive. This characteristic is particularly advantageous for density regressions where zero and negative values signify meaningful states, such as the absence of related activity or negative growth rates. The IHS transformation has been shown to approximate logarithmic transformation for large positive values while retaining the ability to handle zeros and negative values without the need for arbitrary adjustments (Burbidge et al., 1988; MacKinnon & Magee, 1990). IHS transformations are not a solution to the issue of zeros in all cases - the estimated effect from IHS transforms includes both the extensive and intensive margin effects, and with a big mass of zeros present, this can dramatically affect the coefficients and also whether the results appear to be statistically significant (Mullahy & Norton, 2022; Chen & Roth, 2023). However, note that in our case, the growth variable only considers the intensive margin because of the way the growth variable is defined. The extensive margin is considered separately in the "Appearance" regression. This ensures that we have a "two-part" regression, with no mass of zeros, thus addressing the concerns surrounding the IHS transformation.

A.3.4 Cross-domain density

Most countries participate in international trade, a significant proportion have scientific publications, and a relatively smaller group of countries patent. Nevertheless, there might be a significant overlap between required capabilities across domains. We separately identify capability overlaps between scientific fields and technologies and between industries and technologies.

Capability overlaps between scientific fields and technology classes. Recently, Marx & Fuegi (2020) built a database of citations from patents to scientific fields to capture flows of ideas. We aggregate these citations to the technology class - scientific field level. Let's denote the number of citations from technology class t to scientific field f with $C_{t,f}$. We can build a relatedness matrix between each class and field with:

$$\phi_{t,f} = \frac{C_{t,f}}{\sum_{f'} C_{t,f'}}.$$

Given this measure, we can build a density around a technology class based on the scientific output pattern of country c with:

$$D_{ct}^{PS} = \frac{\sum_{f'} M_{cf'}^{S} \phi_{t,f'}}{\sum_{f'} \phi_{t,f'}}.$$

Capability overlaps between industries and technology classes. Firms engage in international trade and also innovate. Using this observation, we can develop a relatedness relation based on firms exporting in industry i and innovating in technology class t. Let's denote the number of firms active in the industry i and technology class t with $N_{t,i}$. We can build a relatedness matrix between each class and industry with:

$$\phi_{t,i} = \frac{N_{t,i}}{\sum_{i'} N_{t,i'}}.$$

Given this measure, we can build a density around a technology class based on the industrial production pattern of country c with:

$$D_{ct}^{PX} = \frac{\sum_{i'} M_{ci'}^{X} \phi_{t,i'}}{\sum_{i'} \phi_{t,i'}}.$$

B Data methods

B.1 Patent data

We rely on patent data compiled by WIPO, which builds primarily on EPO's Patstat 2023a and WIPO's PatentScope datasets. Although our patent data has worldwide coverage of patent offices and inventors' locations, the statistical inference based on raw patent data counts is limited for most developing countries. As shown in the accompanying paper (Hausmann et al., 2024), patent production is highly concentrated in a few countries and limited (or even non-existent) in most of the developing world. These differences in patenting do not necessarily imply that countries with zero patents have no technological capabilities, as patent usage responds to differences in, for instance, industries, firm size, and strategic behaviors (Mezzanotti & Simcoe, 2023). Moreover, there are some concerns about the comparability of patents filed in different patent offices, which limits international comparisons. For these reasons, our analyses:

- Only count international patent families, as defined by Miguelez et al. (2019). That is, "inventions for which the applicant has sought patent protection beyond its home patent office. This definition also includes patent applications by applicants filing only abroad, filing only through the PCT system, or filing only at the EPO" (Miguelez et al., 2019, p. 4).
- Only include countries with over 100 international patent families after the year 2000 (as indicated by their year of earliest filing).
- Only consider countries with a population over 1 million.
- Only include technologies (IPC4 subclasses) that appear in at least ten countries.

B.2 Scientific publication data

Scientific publication data comes from OpenAlex (Priem et al., 2022). We rely on a snapshot of OpenAlex from January 2023. We chose OpenAlex for measures of scientific publications and citation counts due to its open availability, comparability with MAG, and relatively high global coverage (Jiao et al., 2023). OpenAlex was launched in January 2022, following the closure of Microsoft Academic Graph (MAG), which had been an important source of openly available data for bibliometric analyses (Wang et al., 2020). Several studies have validated the usefulness of MAG, and compared it to other existing data sources such as Google Scholar, Scopus, and Web of Science (Wang et al., 2019; Harzing & Alakangas, 2017; Thelwall, 2017; Martín-Martín et al., 2021). Some evidence indicates a general consensus between these databases regarding citation counts, but the criteria for inclusion of papers and classification of document types seem to be inconsistent (Scheidsteger et al., 2023; Jiao et al., 2023). While MAG has lower publication and citation coverage compared to Google Scholar across most disciplines, it has higher coverage as compared to Scopus and Web of Science (Martín-Martín et al., 2021; Harzing & Alakangas, 2017; Thelwall, 2017; Huang et al., 2020). OpenAlex was intended to be MAG's immediate successor and has improved upon MAG's coverage, even for overlapping years of coverage (Scheidsteger & Haunschild, 2022).

Due to the automated nature in which OpenAlex collects information on scientific publications, it does not incorporate the data quality filtering mechanisms used in other databases, such as Scopus. However, it offers a potential advantage in providing broader global coverage. To address potential data quality issues, we implement the following procedure for filtering OpenAlex data:

- Only consider countries with over 5000 publications after the year 2000.
- Only include countries with a population over 1 million.
- Only include scientific fields for a given country in RCA calculations if there are over five publications in that field from that country.
- Only include scientific fields with over ten total publications in 2020.

Additionally, to calculate ECI values, we add a condition that a paper must have received at least five citations to be included in country counts.

It is important to note that we do not apply these filters when analyzing distributions and inequality patterns, as certain countries may have limited publication output, and excluding them could skew our analyses.

B.3 Trade data

We use data on international trade from Bustos & Yildirim (2023), who clean trade data from UN COMTRADE. This cleaning approach improves data quality by accounting for mismatches in reporting by exporters and importers. Furthermore, we only consider countries listed in the Atlas of Economic Complexity and products included in the product space (Hausmann et al., 2014; Harvard's Growth Lab, 2023).

B.4 Clustering countries

In order to cluster countries based on their scientific publication portfolio, we apply the following procedure:

- First, we consider the matrix of countries and their publications in various scientific fields per capita.
- We standardize the values in each scientific field, so that scientific fields are made comparable, ignoring the total quantity of publications in each field and just considering the spread across countries
- We reduce the dimensionality of this data in order to reduce its noise. We rely on UMAP (McInnes et al., 2018), with the parameters n_neighbors set to 500 and min_dist set to 0.1. We set the number of dimensions to 2, so that the reduced data can be easily visualized.
- Crucially, we compute the cosine distance metric while reducing dimensionality. This ensures that any subsequent clustering is due to differences in the "angle" of the publication portfolio and not the magnitude of publications. If we were considering the quantity of publications, countries that publish a lot would cluster together and separately from countries that publish infrequently. Using a distance metric that ignores magnitudes, as in our case, ensures that the effect

we capture indicates the "type" of publications, not the quantity.

- Finally, we cluster countries using a simple K-Means approach. Since the clusters are remarkably separated from each other, results are not sensitive to the type of clustering method used.

We follow a similar approach when demonstrating clustering based on patents. In the case of patents, instead of creating a matrix based on countries and scientific fields, we create one based on countries and IPC 4-digit technology classes. The rest of the clustering approach remains the same.

B.5 Linking patents and scientific publications

The connections between academic fields and technological classes are estimated based on the patent citation links to scientific publications extracted by Marx & Fuegi (2020). In particular, we count the number of patent families in which a technology (i.e., IPC4 subclass) and an academic field (i.e., an OpenAlex concept_id at level 1) appear together. As a patent family can be classified into multiple technologies, cite multiple papers, and those papers themselves are classified into multiple academic fields, we divide every citation link by the total number of IPC4 classes in a patent family, the number of papers cited by a patent family, and the number of concepts in a paper. Hence, every patent family contributes to one unit in the total count, and the total sum equals the number of patent families for which information on IPC4 subclasses is available in Patstat 2021a and which cite at least one paper with at least one concept_id in level one.[11]

B.6 Linking patents and products

The connections between technologies and products are estimated by linking trade data to worldwide patent data. Precisely, we match worldwide exports to Colombia (Colombian

[11] Some patents from Marx & Fuegi (2020) cannot be found in Patstat2021a or do not have any assigned technology. Similarly, there are citations to papers that do not have an assigned concept_id at level one in OpenAlex. Hence, the total sum of considered patent families equals 3,100,191, which corresponds to 89% of all the total patent families initially identified by Marx & Fuegi (2020).

imports worldwide), obtained from Colombia's customs office between 2006 and 2014 (DIAN, 2023), with patents extracted from Google Patents Public Data, from 2000 onwards. This is done by matching exporter names on individual import records to assignee names on patents.[12] From this, it can be inferred, for instance, that firms that patent in semiconductors also tend to export products related to semiconductors more often. Hence, technology codes on patents can be related to product codes by taking those technologies that disproportionally co-occur with products. If a firm exports multiple products, technology counts are weighted by the export share of a product in a firm.

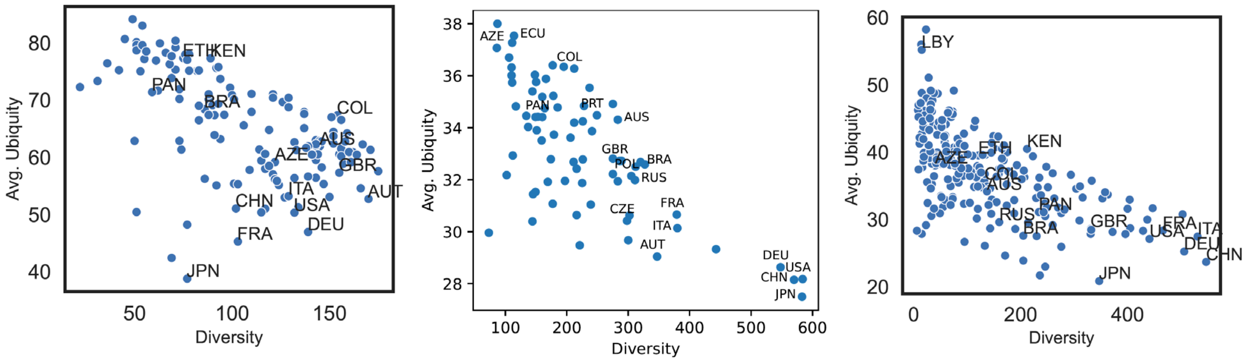
B.7 Other data sources

We use the World Bank's World Development Indicators (WDI) database to source data on population, GDP, R&D spending, and researchers per million people.

[12] We rely on the assignee-name disambiguation (assignee_harmonized variable) provided by Google Patents Public Data, to handle cases in which a single firm is associated with multiple names.

C Additional results

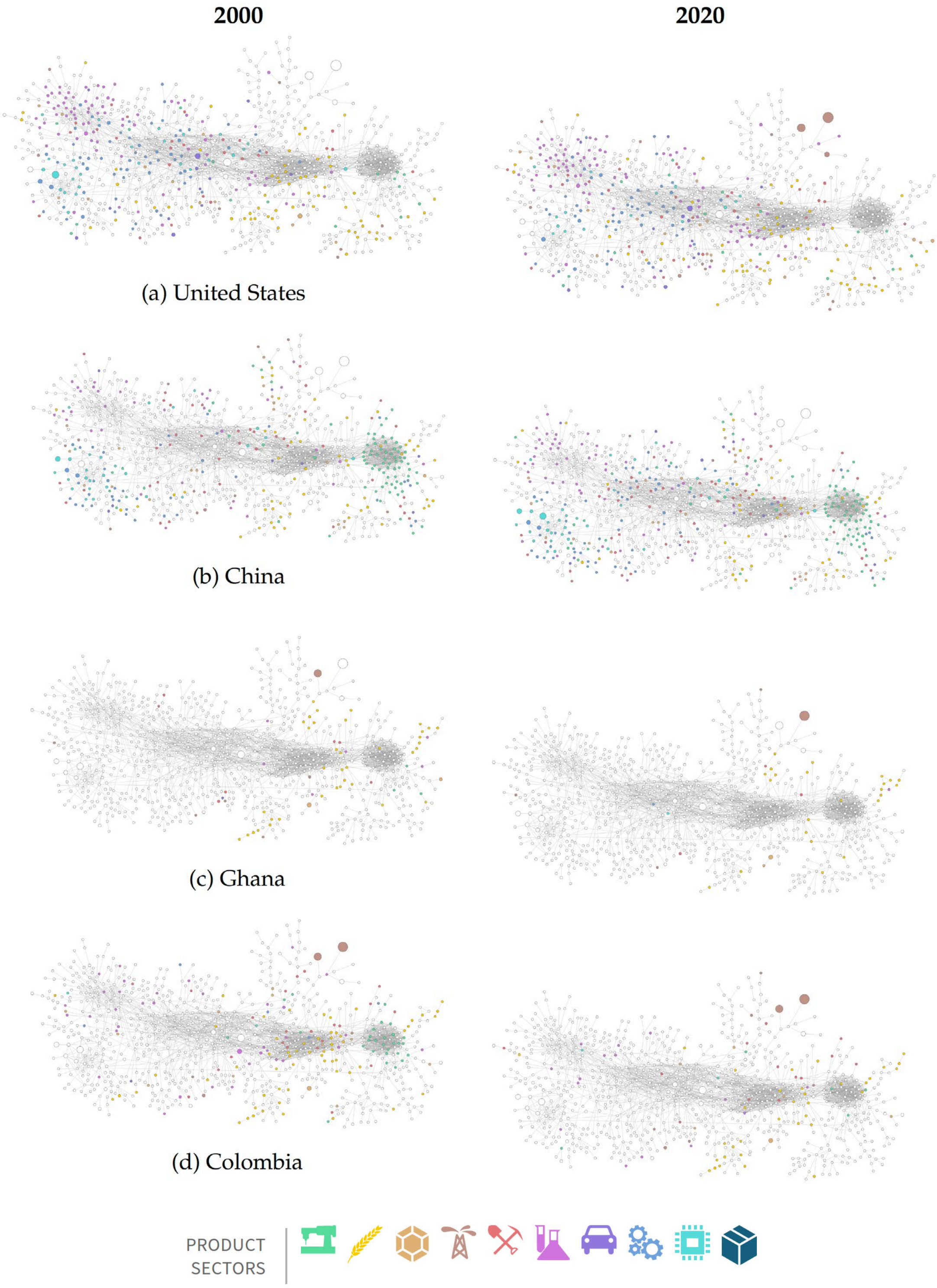
Figure C.1: Diversity and Average Ubiquity, 2015 – 2020



(a) Scientific Publications (b) Patents (c) Trade

Source: Own construction, using trade data from UN COMTRADE, scientific publications from OpenAlex, and patents compiled by WIPO.

Figure C.2: Evolution of the product spaces for selected countries



Source: The Atlas of Economic Complexity (Harvard's Growth Lab, 2023). Note: Nodes (dots) represent products (following the Harmonized System - HS - 1992 classification) and links (lines) their primary connections. Products that are strongly related to one another (i.e., requiring related capabilities) are clustered closer together in the network. Node sizes are based on the product's world trade. Node colors represent the product's major sectors: textiles, agriculture, stone, minerals, metals, chemicals, vehicles, machinery, electronics, and others.

Table C.1: Most important scientific fields for clustering publications-derived country space

Scientific Field	Broad Field
Computational biology	Biology
Meteorology	Environmental science
Bioinformatics	Medicine
Artificial intelligence	Computer science
Classical mechanics	Physics
Cognitive psychology	Psychology
Climatology	Environmental science
Atmospheric sciences	Environmental science
Microbiology	Biology
Polymer chemistry	Chemistry
Statistics	Mathematics
Cell biology	Biology
Genetics	Biology
Biochemistry	Chemistry
Molecular biology	Biology

Table C.2: Most important technology classes for clustering patents-derived country space

IPC4	Subclass	Section
B65D	Containers For Storage Or Transport Of Articles...	Performing Operations Transporting
A61M	Devices For Introducing Media Into, Or Onto, Th...	Human Necessities
B05D	Processes For Applying Liquids Or Other Fluent ...	Performing Operations Transporting
B01D	Separation	Performing Operations Transporting
B23K	Soldering Or Unsoldering Welding Cladding Or Pl...	Performing Operations Transporting
H04W	Wireless Communication Networks	Electricity
C07K	Peptides	Chemistry Metallurgy
F04B	Positive-Displacement Machines For Liquids Pumps	Mechanical Engineering Lighting Heating Weapons...
G07F	Coin-Freed Or Like Apparatus	Physics
C12N	Micro-Organisms Or Enzymes Compositions Thereof...	Chemistry Metallurgy
F24F	Air-Conditioning, Air-Humidification Ventilation...	Mechanical Engineering Lighting Heating Weapons...
B29L	Indexing Scheme Associated With Subclass Relati...	Performing Operations Transporting
G06K	Recognition Of Data Presentation Of Data Record...	Physics
G01F	Measuring Volume, Volume Flow, Mass Flow, Or Li...	Physics
H01B	Cables Conductors Insulators Selection Of Mater...	Electricity

© WIPO, 2024

World Intellectual Property Organization
34, chemin des Colombettes, P.O. Box 18
CH-1211 Geneva 20, Switzerland

Attribution 4.0 International (CC BY 4.0)

This work is licensed under Creative Commons Attribution 4.0 International.

The user is allowed to reproduce, distribute, adapt, translate and publicly perform this publication, including for commercial purposes, without explicit permission, provided that the content is accompanied by an acknowledgement that WIPO is the source and that it is clearly indicated if changes were made to the original content.

Suggested citation:
Hausmann, R., M.A. Yildirim, C. Chacua, S. Gadgin Matha and M. Hartog (2024). Global Trends in Innovation Patterns: A Complexity Approach. WIPO Economic Research Working Paper No. 80. World Intellectual Property Organization.

Adaptation/translation/derivatives should not carry any official emblem or logo, unless they have been approved and validated by WIPO. Please contact us via the WIPO website to obtain permission.

For any derivative work, please include the following disclaimer: "The Secretariat of WIPO assumes no liability or responsibility with regard to the transformation or translation of the original content."

When content published by WIPO, such as images, graphics, trademarks or logos, is attributed to a third-party, the user of such content is solely responsible for clearing the rights with the right holder(s).

To view a copy of this license, please visit https://creativecommons.org/licenses/by/4.0

Any dispute arising under this license that cannot be settled amicably shall be referred to arbitration in accordance with Arbitration Rules of the United Nations Commission on International Trade Law (UNCITRAL) then in force. The parties shall be bound by any arbitration award rendered as a result of such arbitration as the final adjudication of such a dispute.

The designations employed and the presentation of material throughout this publication do not imply the expression of any opinion whatsoever on the part of WIPO concerning the legal status of any country, territory or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

This publication is not intended to reflect the views of the Member States or the WIPO Secretariat.

The mention of specific companies or products of manufacturers does not imply that they are endorsed or recommended by WIPO in preference to others of a similar nature that are not mentioned.

Cover: WIPO Design