European Commission

Divided We Fall Behind

Why a fragmented EU cannot compete in complex technologies

Pierre-Alexandre BALLAND, Valentina DI GIROLAMO, Florence BENOIT, Julien RAVET and Alexandr HOBZA

R&I PAPER SERIES
WORKING PAPER 2025/04

Research and Innovation

Divided We Fall Behind: Why a fragmented EU cannot compete in complex technologies

European Commission
Directorate-General for Research and Innovation
Directorate G — Common Policy Centre
Unit G.1 — Common R&I Strategy & Foresight Service
Contact Julien Ravet
Valentina Di Girolamo
Florence Benoit
Email Julien.RAVET@ec.europa.eu
RTD-ECONOMIC-ANALYSIS@ec.europa.eu
RTD-PUBLICATIONS@ec.europa.eu
European Commission
B-1049 Brussels

Manuscript completed in April 2025

PDF	ISBN 978-92-68-26701-1	doi:10.2777/8548441	KI-01-25-083-EN-N

Luxembourg: Publications Office of the European Union, 2025

EUROPEAN COMMISSION

Divided We Fall Behind

Why a fragmented EU cannot compete in complex technologies

Working Paper

Authors:

Pierre-Alexandre BALLAND, Valentina DI GIROLAMO, Florence BENOIT, Julien RAVET, Alexandr HOBZA

2025 Directorate-General for Research and Innovation

TABLE OF CONTENTS

ABSTRACT

Fragmentation in the European Union's R&I system is increasingly acknowledged but theoretical frameworks and empirical evidence remain scarce. This paper introduces a novel complexity-based approach to analyse the competitiveness costs of R&I fragmentation, focusing specifically on hub connectivity as a key metric. Using a comprehensive dataset combining patent records (OECD RegPat) and scientific publications (OpenAlex) from 2000 to 2023, we examine R&I networks across multiple spatial levels and domain categories. Our analysis reveals three critical findings. First, the European R&I system is much more fragmented compared to the US, with major European hubs showing notably weaker interconnectivity than their US counterparts. Second, we demonstrate that hub connectivity becomes particularly crucial for complex technologies, including AI, biotechnology, and quantum computing. Third, we find that the efficiency gap between the US and Europe is observed across all technology domains but is most pronounced in complex technologies, creating a substantial competitive disadvantage in strategic sectors. These findings have significant implications for European innovation policy and suggest the urgent need for targeted interventions to enhance cross-regional R&I collaboration in complex technological domains.

1. Introduction

We might be experiencing the most rapid and transformative technological revolution in human history. In this era of global competition, the strength of research and innovation systems (RISs) is the cornerstone of economic success and global influence. Despite its considerable collective resources and talent, Europe faces a unique fragmentation challenge. The Draghi report (2024) highlighted significant internal barriers within the EU, which was reiterated in a recent Financial Times column, noting that they effectively impose a 45% trade tariff on manufacturing goods and 110% on services.[1] In this regard, the Letta report (2024) proposed a "fifth freedom" cantered on research, innovation, and education, Meanwhile the European Commission's new Competitiveness Compass calls for the removal of cross-border barriers to enhance competitiveness and strengthen the Single Market.[2]

It is clear that this lack of integration hinders the efficient flow of knowledge and innovation, creating major inefficiencies and missed opportunities for collaboration. It weakens Europe's ability to compete with innovation leaders like the United States (US), and to address global challenges such as climate change and health crises, where cross-border cooperation is essential. Although European leaders have highlighted fragmentation as a key obstacle, the literature has largely failed to provide robust theoretical frameworks or empirical evidence to explain, measure and monitor R&I fragmentation.

This paper addresses this critical gap by introducing a novel complexity-based framework. At its core, our approach examines hub connectivity - the degree to which major R&I centres collaborate and share knowledge. By analysing a comprehensive dataset combining patent applications (RegPat) and scientific publications (OpenAlex) from 2000 to 2023, we provide the first systematic comparison of hub connectivity between Europe and the US innovation systems. Our analysis spans multiple spatial levels (from urban areas and functional urban areas to larger regional units) and covers two domain categories: the aggregate of all technologies (all) and by technological field as classified by the World Intellectual Property Organization (WIPO categories). This multi-scalar approach allows us to capture the multifaceted nature of R&I networks and test the robustness of the framework.

Our analysis highlights three key findings. First, the European R&I system is significantly more fragmented compared to the US, with major European hubs showing notably weaker interconnectivity. Second, we show that hub connectivity is particularly important for complex technologies. Third, fragmentation within Europe is most pronounced in these complex technological domains, posing a major competitive disadvantage in strategic sectors. These findings have significant implications for EU innovation policy and suggest the urgent need for targeted interventions to enhance cross-

---

[1] Forget the US — Europe has successfully put tariffs on itself
[2] Competitiveness compass - Consilium

regional R&I collaboration in more advanced and sophisticated technology fields.

By establishing a clear empirical foundation for understanding R&I fragmentation and its consequences, this paper aims to inform more effective policy interventions to strengthen Europe's position in the global innovation landscape. The findings presented have significant implications for research collaboration, innovation policy, and the future of technological competitiveness in an increasingly complex world.

The remainder of this paper is structured as follows. Section 2 provides an overview of the theoretical framework underpinning the analysis, highlighting the key costs associated with RIS fragmentation, and introducing the hypotheses tested in this paper. Section 3 describes the data employed and develops the proposed empirical framework. Sections 4 and 5 discuss the key results and the robustness checks carried out to validate the empirical findings. Section 6 concludes the paper, discussing the policy implications of the results.

2. The cost of fragmentation

The fragmentation of RISs across Europe significantly undermines its competitiveness in the global innovation landscape. Unlike the US, which benefits from a cohesive and strategically aligned R&I strategy, Europe contends with a multiplicity of approaches that, in the absence of a unifying vision, can weaken its collective potential.

Within Europe, each Member State largely pursues its own innovation priorities, shaped by local economic needs and political agendas. With 27 national innovation strategies, funding for R&I is often dispersed and misaligned across Member States, while only about 10% of total R&I spending is managed through more directed EU-wide programmes (Draghi, 2024). This fragmentation hinders Europe from focusing on a cohesive set of key technologies essential for global competitiveness, leading to suboptimal investment in strategic areas. This disparity is especially problematic for complex technologies, such as AI or quantum computing, which require significant and concentrated investment to reach scale. In contrast, the US benefits from a single national strategy, ensuring a more coordinated allocation of resources and a clearer alignment between investment priorities.

The European R&I fragmentation also leads to massive, missed opportunities for network effects. Scaling complex technologies requires strong network effects, achieved through interconnected ecosystems of researchers, companies, and institutions. Member States' siloed strategies prevent the formation of pan-European networks thereby limiting the cross-border collaboration and knowledge-sharing needed for scaling technologies. The US benefits from a large, unified market supported by federal programmes that actively foster collaboration and knowledge sharing across States. The lack of a coordinated R&I strategy across Member States leads to missed opportunities for collaboration and synergy. This isolation limits the potential for cross-border projects that could leverage diverse expertise and resources, ultimately resulting in a weaker innovation output.

In addition, fragmentation in R&I causes duplicated efforts among Member States, leading to the inefficient use of resources. The lack of coordination often means that multiple entities independently pursue similar projects, which typically result in small-scale efforts that struggle to compete in a technological ecosystem driven by a winner-takes-all model. This wastes both funding and manpower, while also failing to produce meaningful advancements or generate proportional value in terms of knowledge.

The cost of fragmentation is further evident in the incompatibility of policies and standards across Member States. Different regulatory frameworks, labour laws, and visa procedures create friction, making it difficult to deploy and scale advanced technologies across Europe. For instance, a technology developed in Germany may face hurdles in Italy due to differing standards, similarly a non-European software engineer hired in France might struggle to relocate to the Netherlands.

Furthermore, the absence of a true European capital-market union further exacerbates R&I fragmentation. The Draghi report highlights a concerning trend: over one-third of European unicorns relocate to the US, attracted by a more favourable investment climate (Draghi, 2024). This capital flight diminishes the financial resources available for R&I in Europe and signals a loss of confidence in its ability to nurture and support its own innovative enterprises.

Finally, fragmentation in R&I ecosystems makes Europe less attractive to global talent by creating inefficiencies and barriers. Unlike more integrated ecosystems, such as the US, Europe's fragmented approach results in a perceived lack of opportunities and mobility for researchers and innovators. Inconsistent funding mechanisms, regulatory discrepancies, and limited opportunities for cross-border collaboration are huge limiting factors. For example, navigating the French R&I system involves mastering unique rules and networks, creating significant barriers to researcher and firm mobility when transitioning to another European system. In contrast, while the US system also exhibits state-level variations, it offers greater flexibility and facilitates smoother transitions between States and industries as institutional and bureaucratic complexities are more limited (e.g., Shapira, et al., 2010). In Europe, however, targeting one country's system often does not translate into broader mobility or integration across the continent. Over time, this lack of mobility and coordination can exacerbate brain drain, as Europe's most talented individuals seek out ecosystems with deeper resources and easier access to global markets.

2.1 The Geography of R&I collaborative networks in Europe and US

From a historical perspective, the structure of the US innovation economy started to undertake a significant transformation as of the beginning of the 20th century, when an increasing number of collaborative ties between East and West Coasts started to emerge (Abbasiharofteh et al., 2024). Recent evidence also suggests that the US highly innovative hubs tend to collaborate more both nationally and internationally, representing key global innovation hotspots characterised by higher degree of connectivity within the global innovation network (WIPO, 2019). On the contrary, R&I collaborations in Europe still largely occur between actors located within the same national borders, while cross-country collaborations are mostly limited to cross-border regions (European Commission, 2024). Despite the increasing internationalisation of innovation and scientific activities, the integration of the European R&I network still faces important challenges (Chessa et al., 2013), and only a few global innovation hubs have been able to emerge in Europe, notably in the United Kingdom (UK), France and Germany (WIPO, 2019).

As highlighted in the previous section, this gap between the European and the US R&I systems can be attributed to the EU's structure as a union of multiple sovereign States, which makes the European science and innovation space highly heterogeneous. Such heterogeneity translates into a broad set of diverse national priorities that can lead to lack of coordination and to duplication of efforts, with each Member State having its own research agenda, funding mechanisms, and policy frameworks.

Scientific and innovation activities are territorial embedded processes that need specific structural and institutional conditions to thrive (Rodriguez-Pose & Crescenzi, 2008). In a fully integrated research system, collaboration partners should be identified exclusively based on scholarly considerations. However, spatial heterogeneity still represents a non-negligible determining factor, with geographical proximity still playing a key role in facilitating and explaining the participation in collaborative knowledge production activities, especially in relation to the transfer of tacit knowledge (e.g., Okubo & Zitt, 2004; Maggioni & Uberti, 2007; Hoekman et al., 2009; Moreschalchi et al. 2015; Lata et al., 2017). At the same time, other forms of proximities also need to be considered (Boschma, 2005). As an example, institutional differences between countries and regions represent an important challenge to the development of R&I collaborations between different national R&I systems. This is particularly relevant in the European case, where the presence of persistent language, cultural and legal barriers create further obstacles to effective R&I collaborations between European regions and countries, even after controlling for geographical distance (Hoekman et al., 2010; Scherngell & Barber, 2011). The same holds for the US (Singh & Marx, 2013), although recent evidence suggests that the importance of institutional proximity has been diminishing over time (Abbasiharofteh et al., 2024).

Despite the important progress in the economic integration process, key institutional settings in research infrastructures, education systems and labour markets in the EU are still largely defined at national level (Banchoff, 2002; Crescenzi et al., 2007). Additionally, while Europe can count on centralized funding programmes, like the EU Framework Programme (FP) for R&I, fragmentation in resource allocation persists, as a significant portion of R&I funding comes from national or regional sources. This makes the European innovation system still largely dominated by distinct national and regional institutions, with diverse priorities aiming at improving local scientific or technological capacities to boost regional competitiveness (Bristow, 2005; Crescenzi et al., 2007).

Based on these considerations, the first hypothesis that this paper will test is that the European R&I system is more divided and less efficient compared to that of the US:

H1: The R&I system is more fragmented in Europe than in the US

2.2 Complexity and the structure of R&I systems

R&I processes are inherently interactive processes involving an increasing number of interconnected actors (Morgan, 2007; Autant-Bernard et al., 2007). Furthermore, as the knowledge frontier expands, so does the volume of information that needs to be processed to generate new ideas and innovations (Jones, 2009). This increasing "burden of knowledge" adds on the complexity and interconnectedness of the knowledge space, making the invention and innovation process more difficult (Fleming and Sorenson, 2001). This is particularly relevant for more complex technologies, whose development strongly hinges on the availability of multidisciplinary expertise that cannot easily be found in a single location. The diverse and specialised knowledge underpinning this type of technologies is typically less easy to replicate and transfer across regions, making complex technologies particularly "sticky" in space, and explaining why they tend to be deeply rooted in specific environments, where the necessary expertise and resources are concentrated (e.g., Fleming & Sorenson, 2001; Hidalgo & Hausmann, 2009; Balland & Rigby 2017).

In addition, as knowledge and technology clusters grow, they attract more skilled professionals and resources (Marshal, 1920; Porter, 1990). This creates positive self-reinforcing feedback loops, which further increase concentration and the dominance of highly innovative hubs, thereby contributing to the geographic and economic centralisation of specialised and complex knowledge (e.g., WIPO, 2019; Balland et al., 2020). The geographic distribution of complex technologies and scientific activities is, thus, likely to differ from the broader patterns of invention and innovation. Investigating the development over time of breakthrough inventions in the US, Esposito (2023) finds that during the 20th century inventors based in areas endowed with a richer and more diverse pool of ideas were more likely to develop breakthroughs than those located in regions lacking this type of attributes. Since only a few places can provide the right conditions for the necessary competences for the development of complex activities (Hidalgo & Hausmann, 2009), complex knowledge strongly hinges on local interactions and intense collaborations than less complex one (e.g., Balland & Rigby 2017; Van der Wouden, 2020; Abbasiharofteh et al., 2024).

As noted by Balland & Rigby (2017), the development of complex technologies strongly relies on the transfer of tacit knowledge, which binds them to specific geographic regions. However, as the knowledge space becomes more complex, important innovation locations cannot solely rely on the knowledge inputs coming from their countries' national science and innovation systems, but rather need to increase their participation in international collaboration networks to facilitate the exchange of ideas and skills across borders (WIPO, 2019). As specialised knowledge clusters, accessing expertise that are outside a given area's core strengths calls for the creation of new bridges across both knowledge fields and geographic locations, thereby making the global integration of collaborative efforts essential for complex innovation and scientific activities to thrive (Cantwell & Salmon, 2018).

Understanding how complexity can mediate the relationship between collaborations and geography is receiving increasing attention from both the academic and policy literature. Recent empirical evidence suggests the existence of a strong positive relationship between complexity and inventors' collaborations across regions and countries in both Europe and the US (Van der Wouden, 2020; European Commission, 2024). Notwithstanding that the "local buzz" within innovation clusters remains essential, this evidence also suggests that strengthening international cooperation, knowledge exchange, and integration of R&I systems is crucial to fostering knowledge creation and increasing the overall competitiveness of R&I ecosystems, particularly in more complex fields (Bathelt et al., 2004).

To summarise, we expect complex technologies to suffer disproportionately from fragmentation because their development involves multiple interdependent components and knowledge domains. Based on this consideration, this paper will contribute to the literature by testing the following second hypothesis:

H2: Integration in the R&I system is particularly important for complex scientific fields and technologies.

Additionally, given the aforementioned bottlenecks faced by the European R&I system, as well as the important innovation gap between Europe and the US (European Commission, 2024; Draghi, 2024), we also postulate another hypothesis. We consider that the impact of fragmentation on technological development can be understood through a multiplicative effect. Both Europe and the US face a baseline level of fragmentation that similarly affects simpler technologies. However, for complex technologies, fragmentation acts as a multiplier, amplifying coordination challenges due to their exponentially greater need for interactions and knowledge transfers across diverse domains. In Europe, institutional, linguistic, and regulatory barriers compound these challenges far more than in the US, where a more integrated system mitigates such effects. As a result, while the Europe-US gap is modest for simpler technologies, it becomes significantly larger for complex technologies, where fragmentation hits hardest. Therefore, we propose to test the following third hypothesis:

H3: The degree of R&I fragmentation in Europe (compared to the US) is particularly pronounced in complex technologies.

3. Data, measures and methods

3.1 Sample and regional definitions

We test our hypotheses using two separate datasets, each containing information on collaborations within Europe and the US at urban area level, respectively. The choice of focusing on the European network including the 27 Member States, the UK and the EFTA countries is based on several considerations. First, given the time span considered in the analysis, the inclusion of the UK ensures consistency between pre- and post-Brexit periods. The UK's recent association to Horizon Europe further justifies its inclusion. Second, both the UK and EFTA countries are important contributors to European R&I. As the EU actively promotes collaborations with these nations to strengthen the European Research Area (ERA), excluding them would overlook key nodes of the European collaborative network and potentially distort the overall picture of European R&I activity. Third, ensuring a broader understanding of European R&I collaborative activity when benchmarking it with the US remains key also from a policy perspective, allowing the analysis to deliver more comprehensive insights to inform strategic decisions.

To define urban areas, we follow the classification provided by the United Nations as published in the World Urbanization Prospect report (see United Nations, 2018). This classification largely relies on national definitions, reflecting the definitions and criteria established by national authorities. The analysis covers 168 urban areas within Europe and 144 urban areas in the US, between 2000-2023. For publications, the sample includes 166 urban areas in Europe and 142 in the US.

3.1.1. Patent data

To collect information on patents, OECD REGPAT (release version February 2024) was used. Patents contain information on the address of inventors and assignees, citations, and information on the technical content (IPC technology classes), amongst others. We use information on patent applications filed under the Patent Cooperation Treaty (PCT) and geocoded addresses of inventors listed on each patent to assign patents to urban areas (Cresenzi et al., 2016). We identify patent collaborations between urban areas based on the collaborations between co-inventors listed on the same patent.

Collaborations are further categorized by technology domain. Patents were assigned to 36 technological fields using the WIPO classification system following Schmoch (2008), which maps IPC codes to broader technology groups.

3.1.2. *Publication data*

Publication data was sourced from OpenAlex, a comprehensive open-source bibliographic database encompassing over 209 million publications across 221 countries and territories. Publications contain information on the institutional affiliation of the author, which was geocoded to link each publication to a particular urban area. We identify publication collaborations between urban areas based on the collaborations between co-authors listed on the same publication. Publications were categorised into scientific domains for further analysis.

To classify individual research topics from OpenAlex into broader categories, we employed a multi-step methodology combining advanced embedding techniques, statistical analysis, and LLM-based decision-making. First, we extracted embedding vectors for both research topics and broader categories using OpenAI's "text-embedding-3-large" model, which captures high-dimensional semantic representations. To establish initial associations, we computed embedding similarity scores between topics and categories, leveraging cosine similarity to quantify semantic proximity. Next, we refined these associations by calculating embedding relatedness density, which measures the embedding similarity of co-occurring topics within the dataset. This step allowed us to assess the actual contextual and relational connections between topics, moving beyond purely semantic similarity to capture real-world co-occurrence patterns. Based on these refined associations, we used OpenAI API (GPT-4o) to perform last step categorization, where the LLM integrated the embedding similarity and relatedness density metrics to make informed classification decisions. Finally, to minimize false positives, we manually defined a threshold for inclusion, ensuring that only robust and meaningful topic-category mappings were retained. This hybrid approach combines the strengths of embedding-based semantic analysis, statistical validation, and LLM reasoning to achieve a nuanced, scalable and accurate classification of research topics.

3.2 The research and innovation system gap

To quantify the fragmentation and efficiency of the RISs, the paper introduces a new indicator, namely the Research & Innovation System Gap (RISG). It measures the correlation between the actual observed and expected connections among spatial units, such as urban areas.

Observed connections are the actual collaborative interactions or links (edges) between two spatial units (nodes), such as urban areas or functional urban areas (FUAs), derived from co-inventor locations listed on patents or publication documents. For example, in Figure 1 we have 5 inventors; 2 of them are co-located in Santa Clara, California (CA). In this case we will have one link between Santa Clara and London, and also a link within Santa Clara, CA. But we do not double count the two links between London and Santa Clara, CA.

Figure 1. Counting links

US 20240265690A1

(19) **United States**
(12) **Patent Application Publication** (10) **Pub. No.: US 2024/0265690 A1**
Anandkumar et al. (43) **Pub. Date: Aug. 8, 2024**

(54) **VISION-LANGUAGE MODEL WITH AN ENSEMBLE OF EXPERTS**

(71) Applicant: **NVIDIA Corporation**, Santa Clara, CA (US)

(72) Inventors: **Animashree Anandkumar**, Pasadena, CA (US); **Linxi Fan**, Santa Clara, CA (US); **Zhiding Yu**, Santa Clara, CA (US); **Chaowei Xiao**, Tempe, AZ (US); **Shikun Liu**, London (GB)

(21) Appl. No.: **18/544,840**

(22) Filed: **Dec. 19, 2023**

Related U.S. Application Data

(60) Provisional application No. 63/483,696, filed on Feb. 7, 2023.

Publication Classification

(51) **Int. Cl.**
G06V 10/82 (2006.01)
G06V 10/80 (2006.01)

(52) **U.S. Cl.**
CPC ***G06V 10/82*** (2022.01); ***G06V 10/811*** (2022.01)

(57) **ABSTRACT**

A vision-language model learns skills and domain knowledge via distinct and separate task-specific neural networks, referred to as experts. Each expert is independently optimized for a specific task, facilitating the use of domain-specific data and architectures that are not feasible with a single large neural network trained for multiple tasks. The vision-language model implemented as an ensemble of pre-trained experts and is more efficiently trained compared with the single large neural network. During training, the vision-language model integrates specialized skills and domain knowledge, rather than trying to simultaneously learn multiple tasks, resulting in effective multi-modal learning.

Expected connections represent how collaborations should ideally be distributed in an optimal, fully integrated system where exchanges occur in proportion to the size and potential of each region. To measure the expected number of connections, we first calculate the theoretical maximum number of links between regions based on their degree of centrality (i.e., the number of direct connections a node has). For example, if Santa Clara, CA has 100 connections and London has 80, the maximum possible links between them would be the average of these values, 90.

We then compute the expected number of connections between regions by determining the fraction of the total theoretical maximum that each pair of regions represents, and weigh this by the total observed links in the network. This ensures that the sum of expected links matches the total number of observed links. Essentially, this method distributes the total observed links across regions in proportion to their respective size thereby giving an expectation of how many links each pair should have if the observed links were distributed according to the theoretical maximum proportions (i.e. based on the sum of their weighted degree centrality) (Granovetter, 1985).

We then calculate the RISG as the correlation coefficient between the observed and expected links at the network level. This correlation serves as an indicator of integration and efficiency within a R&I system. A low RISG indicates that actual connections fall short of what would be expected, which signals network inefficiency and major system failure. This is in line with principles of social network analysis and the gravity model of collaboration in network science: just as larger masses in physics exert stronger gravitational pulls, larger urban centres are expected to attract and form more connections within R&I networks. When major cities are not as well-connected as expected, it suggests the presence of barriers or inefficiencies, rather than a well-integrated system where knowledge, resources, and innovation flow as freely as they should, reflecting the need for Letta's 5th freedom (Letta, 2024).

In contrast, a higher RISG suggests less fragmentation and inefficiencies as the actual network structure aligns more closely with what would be expected in a perfectly integrated system. This suggests that knowledge and ideas can flow more freely across the R&I system and that R&I activities are occurring where they are most likely to be productive, based on the size and capacity of each region. Generally, the closer the network is to the expected distribution of links, the more efficient the system is likely to be as a whole. This efficiency can translate into faster innovation, better use of resources, and potentially higher economic returns from R&I activities.

In essence, the RISG provides a quantitative measure of how close a R&I system is to an optimal and fully integrated network. It can be applied to any general set of collaborations as well as to a specific set of technologies. By comparing observed and expected collaborations between regions, RISG allows for meaningful comparisons across different systems, enabling policymakers to evaluate the current state of integration within their RISs.

While in this paper, the RISG is implemented using patent and publication data, it can be used in a much broader context to measure various network aspects such as funding or mobility. Understanding these gaps can help policymakers and researchers identify where interventions might be needed.

3.3 Poisson pseudo-maximum likelihood (PPML): model and variables

To show the robustness of the results, the paper also includes a controlled version of the RISG using econometric estimations. Given that the analysis examines collaboration patterns between urban areas by counting their connections, the likely presence of overdispersion and a large number of zero observations makes quasi-Poisson regression, negative binomial regression, or zero-inflated models the preferred approaches. Specifically, a Poisson Pseudo-Maximum Likelihood (PPML) is chosen over Negative Binomial models for estimating a gravity model since the latter underperforms when the number of observed zeros exceeds the number of zeros predicted by the model, e.g., in the case of a complete lack of co-patenting between regions due to reasons other than distances and differences in preferences and specializations (Burger et al., 2007). Additionally, the PPML estimator offers a more tractable empirical strategy for incorporating fixed effects and handling multilateral resistance terms[3], making it a robust choice for estimating collaboration intensity through co-patents and co-publications between geographical units (Santos-Silva & Tenreyo, 2006; Fally, 2015; Santos-Silva & Tenreyo, 2022).

To assess whether the R&I system is more fragmented in Europe compared to the US (Hypothesis 1), we specify the model as follows:

$$(1)\ GeoLinks_{i,j,t} = \alpha + \beta_1 Mass_{i,j,t} + \beta_2 Dist_{i,j} + \beta_3 SameCountry_{i,j} + \omega_t + \varepsilon_{i,j,t}$$

where $GeoLinks_{i,j,t}$ is a count variable indicating the number of (patent or publication) collaboration links between urban area i and urban area j, for a given technological or scientific domain c, in year t. As previously mentioned, a collaboration link is created when two co-inventors/authors are named on the same patent/publication with both located in a different urban area.

$Mass_{i,j,t}$ captures the combined size of the two interacting urban areas proxied by the total number of collaborations of urban areas i and j observed in year t. $Dist_{i,j}$ refers to the geographical distance between the two urban areas, measured as the great-circle distance between the centres of the two urban areas (e.g., Hoekman et al., 2010; Lata et al., 2017). The term ω_t denotes the time fixed effects, introduced to account for unobserved time-specific factors, whereas $\varepsilon_{i,j,t}$ denotes the error terms. Since observations in pairs of regions per technology/scientific field are likely to be dependent across years, robust standard errors are clustered to control for error correlation in the panel (Montobbio & Sterzi, 2013).

Specification (1) is extended to examine how integration in the R&I system affects complex technologies and scientific fields (Hypothesis 2), as well as to assess the extent to which R&I fragmentation in Europe is particularly pronounced in complex technologies compared to the US (Hypothesis 3):

$$\begin{aligned}(2)\ GeoLinks_{i,j,c,t} &= \alpha + \beta_1 Mass_{i,j,c,t} + \beta_2 Dist_{i,j} + \beta_3 TechProximity_{i,j,c,t} \\ &\quad + \beta_4 SameCountry_{i,j} + \omega_c + \omega_t + \varepsilon_{i,j,t}\end{aligned}$$

The additional control $TechProximity_{i,j,c,t}$ captures the technological proximity between i and j. Technological proximity estimates the difference in presence

[3] In a standard gravity model with trade data, multilateral resistance terms capture the barriers to trade that each country faces with all its trading barriers.

of knowledge capabilities and absorptive capacity[4] in both locations. If the technological proximity between actors is too low, it may discourage collaborations as both actors may not be able to understand and exploit new complementary knowledge (e.g., Boschma, 2005; Cohen & Levinthal, 1990). $SameCountry_{i,j}$ denotes a dummy variable taking value 1 if the two urban areas are located within the same country or State for Europe and the US, respectively. This variable approximates institutional proximity as being located within the same country, sharing the same language and/or similar values, norms, routines and formal regulations can facilitate collaboration (Maskell & Malmberg, 1999; Boschma, 2005). The two terms ω_c and ω_t represent respectively the technology/scientific domain c fixed effect and the time t fixed effect. As before, robust standard errors are clustered.

All continuous variables were logarithmically transformed. The logarithmic transformation in the context of our model allows the estimated coefficients to be interpreted as elasticities. To ease the comparison between the results obtained for Europe and the US, we standardise the continuous variables included in our model. Descriptive statistics for all the variables are presented in Table 1 and their correlation coefficients are presented in Table 2. Both tables reflect the variables in the sample with aggregated domains.

Table 1 indicates that collaboration links are more frequent for publications than patents, with the US showing a higher mean for patent links, suggesting a more interconnected innovation network. The mass variable is larger in the US, particularly for patents, potentially reflecting stronger innovation hubs. Geographical distance between collaborating urban areas is greater in the US for both patents and publications, consistent with its larger land area and the possibility that long-distance collaborations are more common. Meanwhile, the same country/state variable shows that intra-country collaborations are slightly more prevalent in the Europe for patents and in the US for publications.

Table 1. Descriptive statistics

	Patents			
	Europe		United States	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
$GeoLinks_{i,j,t}$	1,57	13,38	8,70	49,62
$Mass_{i,j,t}$	837,59	929,13	3086,75	3369,32
$Dist_{i,j}$	1101,58	622,89	1861,51	1244,81
$SameCountry_{i,j}$/State	0,10	0,31	0,05	0,21

[4] Knowledge capabilities refer to the ability to create, access and apply knowledge effectively while absorptive capacity is the ability to recognize, assimilate and apply external knowledge.

	Publications			
	Europe		United States	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
$GeoLinks_{i,j,t}$	215,21	926,28	160,61	767,27
$Mass_{i,j,t}$	100120,00	153838,50	90250,08	119040,60
$Dist_{i,j}$	1851,51	1238,02	1107,47	626,30
$SameCountry_{i,j}$/State	0,04	0,21	0,10	0,30

Note: The descriptive statistics are non-standardized and non-log transformed

Table 2 indicates a positive correlation between mass and geo links, which is stronger in the US for patents and publications than in Europe. Distance negatively correlates with geo links in both cases, suggesting that proximity facilitates collaboration. However, the effect is more pronounced in Europe for patents compared to the US, suggesting that geographic constraints may be stronger in Europe. The correlation between being in the same country or state and collaboration is positive across both regions, though its influence is slightly stronger in Europe for patents than in the US.

Table 2. Correlations

		Patents							
		Europe				United States			
	Variable	1	2	3	4	1	2	3	4
1	$GeoLinks_{i,j,t}$	1				1			
2	$Mass_{i,j,t}$	0,2618	1			0,2653	1		
3	$Dist_{i,j}$	-0,1408	-0,2134	1		-0,0558	0,0719	1	
4	$SameCountry_{i,j}$/State	0,2424	0,0168	-0,4065	1	0,1309	0,0316	-0,2673	1

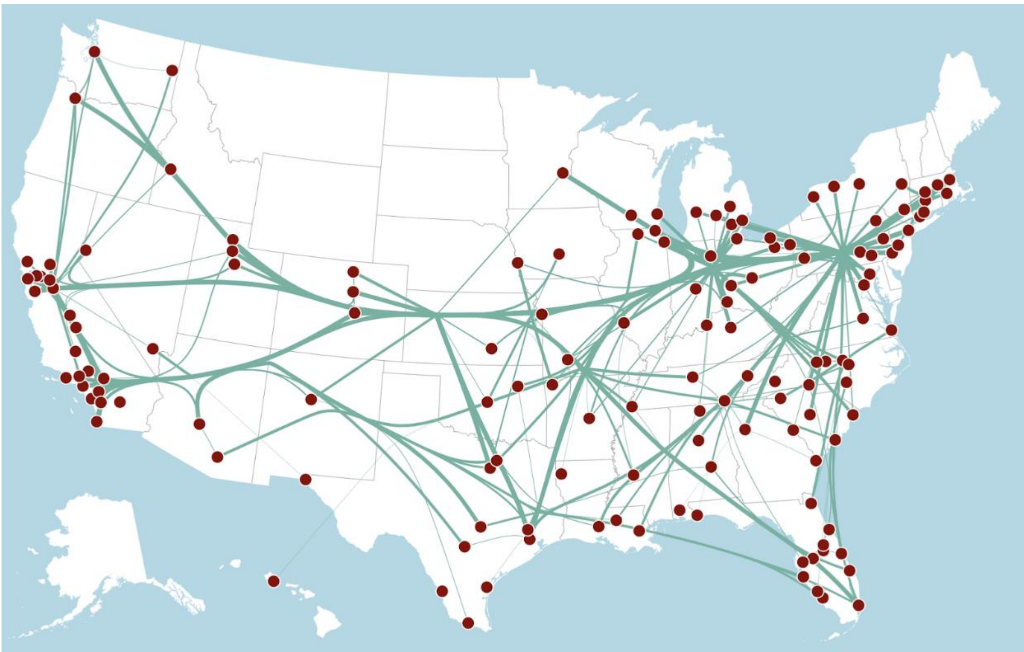
		Publications							
		Europe				United States			
	Variable	1	2	3	4	1	2	3	4
1	$GeoLinks_{i,j,t}$	1				1			
2	$Mass_{i,j,t}$	0,4710	1			0,6223	1		
3	$Dist_{i,j}$	-0,1214	-0,0606	1		-0,0263	0,0095	1	
4	$SameCountry_{i,j}$/State	0,2210	-0,0116	-0,4049	1	0,0109	-0,0212	-0,2716	1

4. Results and discussion

4.1 The European R&I system is more fragmented than the US

When looking at patents, the US demonstrates a more cohesive R&I system compared to Europe, confirming Hypothesis 1. The US system exhibits a strong RISG[5] with correlation of 0.68 between observed and expected links, indicating that actual collaborative relationships between research hubs largely align with predictions based on their size and capacity. This is visually confirmed through network analysis[6] showing connections that span across the entire country without significant border-related clustering (Figure 2, denser lines indicate a larger number of links).

Figure 2. R&I systems in the US (patents)



Note: For an interactive visual representation, you can visit the following webpage: https://www.paballand.com/asg/network-complexity/complexity-regpat-pct-2019-2023-us.html

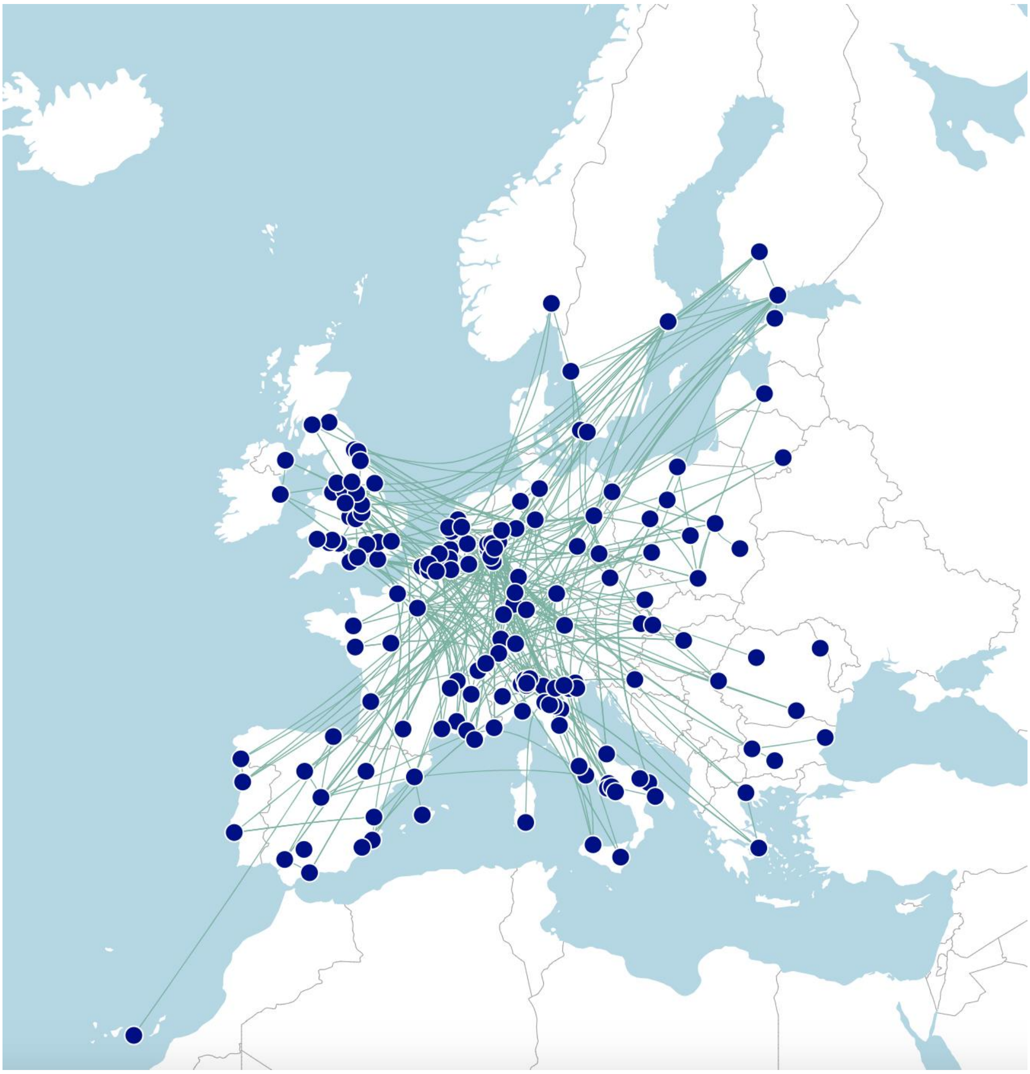
---

[5] We transformed both the expected and observed link counts using logarithmic scaling and excluded all internal connections from the analysis. When we performed alternative analyses using rank correlation coefficients and untransformed metrics, while also including internal connections, we obtained qualitatively similar results.

[6] Huge thanks to Carlos Navarrete and Francisco Rios for their invaluable contributions in developing the front-end for our dynamic network visualisation, including seamless interactivity, real-time rendering, and edge bundling techniques.

In contrast, the European R&I system shows considerable fragmentation, with a markedly lower RISG correlation of 0.40 for patent data. The European network displays distinct clustering around national boundaries, suggesting that country borders still act as significant barriers to R&I collaboration (Figure 3). This substantial gap between expected and observed connections indicates that factors beyond size and distance, such as national policies, language barriers, or institutional frameworks, play a more significant role in shaping European R&I partnerships.

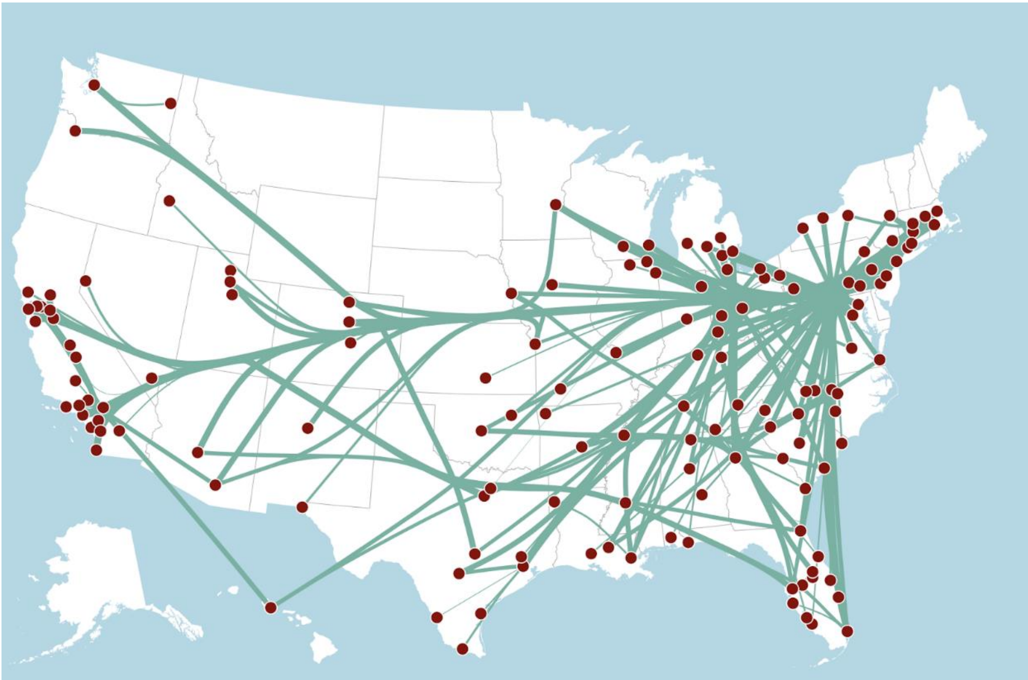
Figure 3. R&I systems in Europe (patents)



Note: For an interactive visual representation, you can visit the following webpage: https://www.paballand.com/asg/network-complexity/complexity-regpat-pct-2019-2023-eu.html

The scientific publication landscape reveals a somewhat different picture, though still maintaining the US-Europe disparity. The US achieves an impressive RISG of 0.81 for scientific publications, indicating an extremely well-integrated academic research network. This suggests that US academic institutions collaborate effectively across state boundaries, creating a more unified research ecosystem (Figure 4).

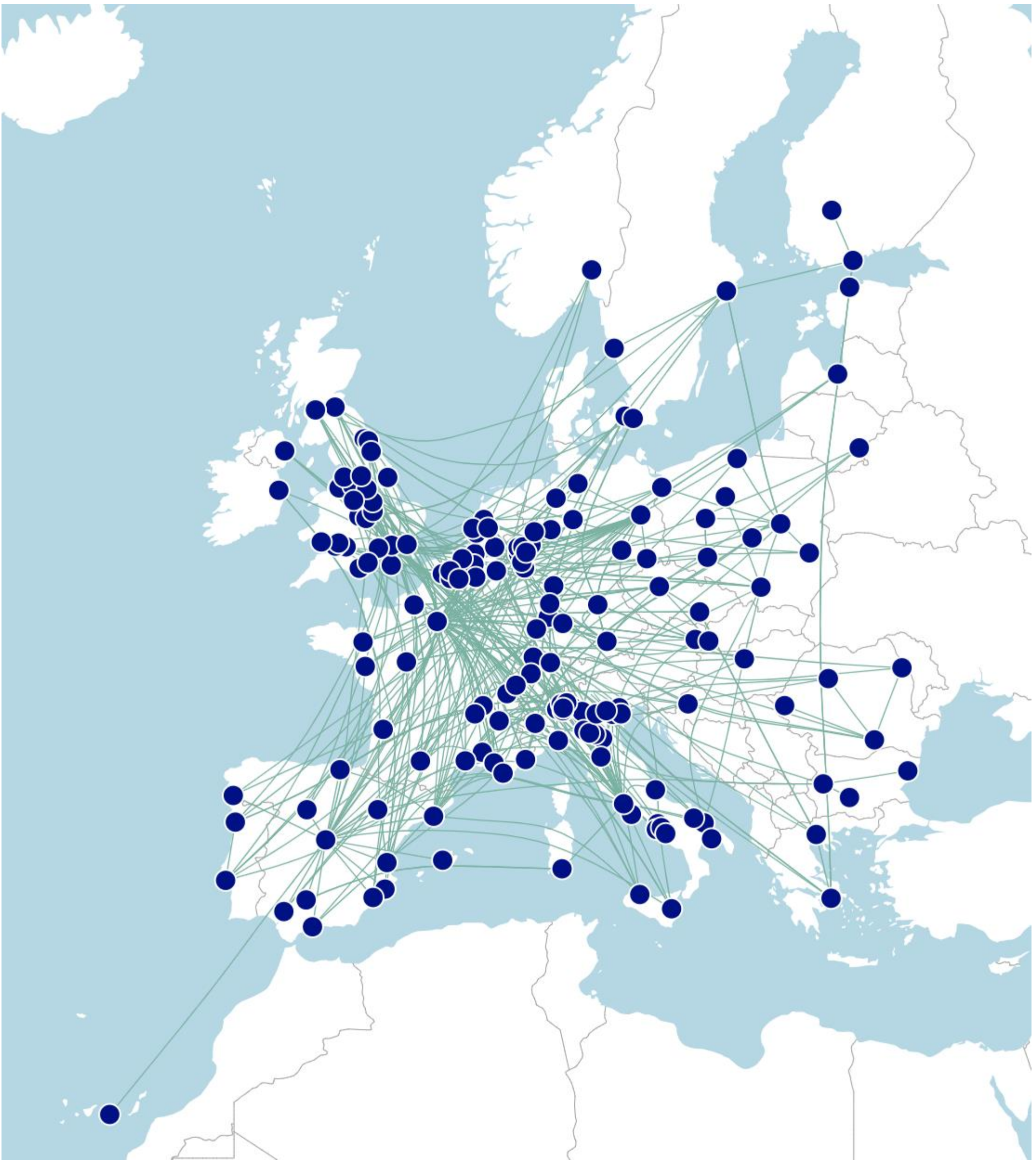
Figure 4. R&I systems in the US (publications)



Note: For an interactive visual representation, you can visit the following webpage: https://www.paballand.com/asg/network-complexity/complexity-openalex-2019-2023-us.html

While Europe shows stronger integration in scientific publications compared to patents, with an RISG of 0.76, it still lags behind the US benchmark. This higher RISG for European publications versus patents (0.76 compared to 0.40) suggests that academic research in Europe has achieved greater cross-border integration than industrial R&D activities. This could potentially be due to structural mechanisms fostering cross-border academic collaborations, such as EU-wide funding frameworks, shared data and open science norms. However, the persistent gap between EU and US figures indicates that even in academic research, Europe has not yet achieved the same level of integration as the US. This is despite having hubs spatially closer in Europe than in the US (Figure 5).

Figure 5. R&I systems in Europe (publications)



Note: For an interactive visual representation, you can visit the following webpage: https://www.paballand.com/asg/network-complexity/complexity-openalex-2019-2023-eu.html

To validate the findings of the RISG analysis, we employ a PPML regression model incorporating high-dimensional fixed effects to examine the fragmentation of R&I systems in Europe and the US. Our approach to model specification involves a stepwise inclusion of the regressors considered, allowing us to evaluate the incremental contribution and statistical significance of each variable. Table 3 to 6 report the results from estimating specification (1) for patent and publication data, respectively.

Table 3. Econometric results for patent collaborations in Europe

	(1)	(2)	(3)	(4)
$Mass_{i,j,t}$	2.405***	1.928***	1.868***	1.799***
	(0.117)	(0.061)	(0.065)	(0.049)
$Dist_{i,j}$		-0.901***		-0.554***
		(0.018)		(0.026)
$SameCountry_{i,j}$			2.981***	1.852***
			(0.062)	(0.083)
Time Fixed Effects	Yes	Yes	Yes	Yes
No. Observations	68642	68642	68642	68642
Wald Chi-sq	424.591	4179.690	2734.138	5329.227

Note: The continuous explanatory variables are taken in natural logarithms; their coefficients can be interpreted as elasticities. *, **, *** indicate significance at 10%, 5% and 1 %. Cluster-robust standard errors are shown in parentheses. When limiting the sample to the 27 Member States of the European Union, similar results were found for Mass and Distance while Same Country coefficients were even higher.

Table 4. Econometric results for patent collaborations in the US

	(1)	(2)	(3)	(4)
$Mass_{i,j,t}$	1.989***	2.009***	1.904***	1.997***
	(0.078)	(0.057)	(0.061)	(0.054)
$Dist_{i,j}$		-0.494***		-0.472***
		(0.031)		(0.026)
$SameState_{i,j}$			1.303***	0.132
			(0.134)	(0.083)
Time Fixed Effects	Yes	Yes	Yes	Yes
No. Observations	51480	51480	51480	51480
Wald Chi-sq	655.681	1282.790	985.844	1551.207

Note: The continuous explanatory variables are taken in natural logarithms; their coefficients can be interpreted as elasticities. *, **, *** indicate significance at 10%, 5% and 1 %. Cluster-robust standard errors are shown in parentheses.

As concerns patent collaborations, the results displayed in Tables 3 and 4 confirm that the size of the interacting urban areas positively influences the likelihood for collaborations. With the exception of the simplest specification estimated, the US consistently reports slightly higher and statistically significant mass ($Mass_{i,j,t}$) coefficients compared to Europe, reflecting a stronger role of urban hubs as anchors for national-level R&I activities. The observed negative effect of geographical distance ($Dist_{i,j}$) in both areas is in line with the existing literature. This effect is notably stronger within Europe, indicating the greater role of geographical distance as a barrier to collaborations. In contrast, the smaller coefficient for the US can be interpreted as a further indication of a more integrated system that facilitates collaborations across greater distances.

Another notable finding is the strong positive effect of being located in the same country ($SameCountry_{i,j}$ for Europe) or same State ($SameState_{i,j}$ for US). The stark difference between both coefficients highlights the dominance of intra-country collaborations in Europe, pointing to the existence of structural barriers (e.g., language barriers and regulatory differences), that impede cross-border interactions. By contrast, the US is characterised by a more balanced pattern of collaboration across States' boundaries, reflecting a more integrated innovation system.

Table 5. Econometric results for scientific collaborations in Europe

	(1)	(2)	(3)	(4)
$Mass_{i,j,t}$	2.313***	2.294***	2.272***	2.281***
	(0.052)	(0.051)	(0.034)	(0.036)
$Dist_{i,j}$		-0.519***		-0.131***
		(0.014)		(0.019)
$SameCountry_{i,j}$			1.884***	1.671***
			(0.037)	(0.055)
Time Fixed Effects	Yes	Yes	Yes	Yes
No. Observations	68310	68310	68310	68310
Wald Chi-sq	1966.521	2790.104	4812.953	4752.349

Note: The continuous explanatory variables are taken in natural logarithms; their coefficients can be interpreted as elasticities. *, **, *** indicate significance at 10%, 5% and 1 %. Cluster-robust standard errors are shown in parentheses. Similar results are obtained for patent collaborations taking into account the 27 Member States of the European Union.

Table 6. Econometric results for scientific collaborations in the US

	(1)	(2)	(3)	(4)
$Mass_{i,j,t}$	2.911***	2.891***	2.925***	2.898***
	(0.052)	(0.049)	(0.052)	(0.049)
$Dist_{i,j}$		-0.165***		-0.153***
		(0.020)		(0.023)
$SameState_{i,j}$			0.504***	0.155
			(0.083)	(0.101)
Time Fixed Effects	Yes	Yes	Yes	Yes
No. Observations	49914	49914	49914	49914
Wald Chi-sq	3171.182	3621.697	3230.525	3607.233

Note: The continuous explanatory variables are taken in natural logarithms; their coefficients can be interpreted as elasticities. *, **, *** indicate significance at 10%, 5% and 1 %. Cluster-robust standard errors are shown in parentheses.

In terms of scientific collaborations, the insights from the RISG analysis are corroborated by the econometric results (Tables 6 and 7). Although the US continues to report higher coefficients for mass ($Mass_{i,j,t}$), the gap with Europe is less pronounced than for patent collaborations. This result further supports our finding that Europe has overall made more progress in integrating its research system as compared to its innovation network. Nevertheless, the consistently positive and statistically significant coefficients associated with being located in the same country ($SameCountry_{i,j}$) confirm that Europe remains bounded by national borders and institutional constraints, limiting its capacity to further strengthen cross-border scientific collaborations.

4.2 R&I systems of complex technologies are less fragmented and more efficient

In order to validate hypothesis 2, we examine the relationship between technological complexity[7] and the mass coefficients derived from separate regression analyses for each technology category[8]. The mass coefficient (as

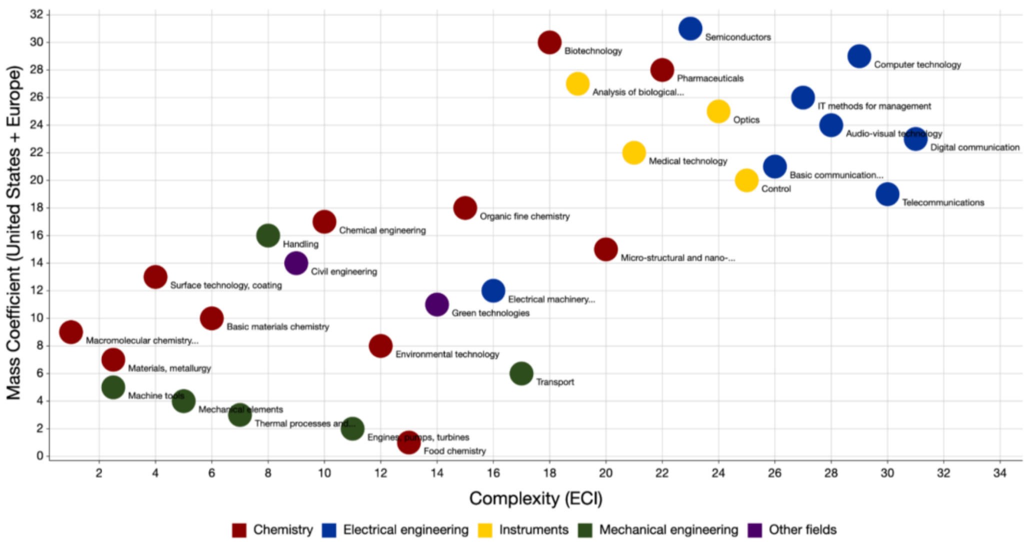
[7] To measure complexity, patent applications were assessed based on technological diversity (i.e., the range of technologies an economy specialises in) and ubiquity (i.e., the number of economies specialising in a given technology) following Hidalgo & Hausman (2009) and Balland & Rigby (2017). Higher values for the knowledge complexity index (KCI) indicate greater technological diversity and specialization in less common technologies, reflecting deeper knowledge and capabilities.

[8] Specifically, the coefficient estimates are obtained by estimating specification (2) as outlined in Section 5.3 using data on the entire US-Europe network (i.e., factoring in also US-Europe collaborations).

defined in Section 3.3) serves as our key metric for measuring R&I efficiency within specific technological domains.

Figure 6, based on patent data, shows a clear positive (rank) correlation between mass and complexity. At the lower end of the complexity spectrum, we find more traditional technologies such as machine tools, textile and paper machines, as well as materials and metallurgy for which hubs are expected to be less connected. In contrast, the upper right quadrant of the graph is populated by highly complex technologies, such as computer technology, digital communication, medical technology, and pharmaceuticals. These technologies typically have the highest levels of R&I efficiency, implying strong collaborative networks and effective knowledge exchange systems.

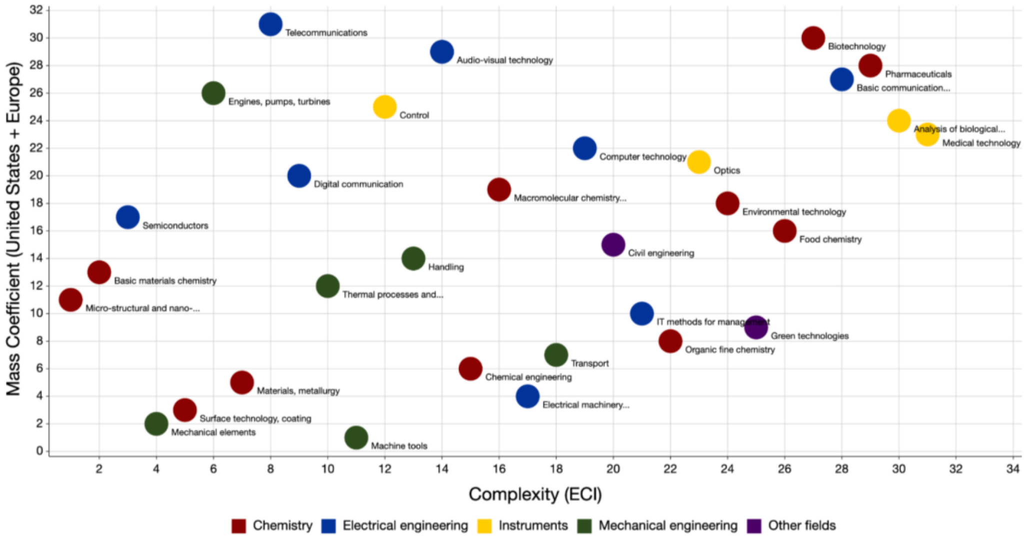
Figure 6. Complexity and R&I efficiency (patents)



A similar pattern is observed for publication data (Figure 7). These results support the notion that more complex technologies not only benefit from but also require greater integration and collaboration in their R&I processes, thereby validating Hypothesis 2.

These results highlight the importance of fostering pan-European collaborative efforts, especially in advanced technological fields, to maximize innovation potential and overcome barriers to knowledge exchange. Additionally, our results also highlight areas where increased collaborative efforts could yield significant benefits (particularly in highly complex technological domains) and the need for targeted strategies to enhance R&I efficiency in sectors that currently underperform relative to their complexity level.

Figure 7. Complexity and R&I efficiency (publications)



4.3 The fragmentation gap is more pronounced for complex technologies in Europe

In order to validate hypothesis 3, we conduct separate regression analyses for the US and Europe, extracting mass coefficients for each WIPO-defined technology category. This comparative approach allows us to directly assess how R&I efficiency differs between the two regions across various technological domains.

Figure 8 reveals a compelling pattern in the comparative efficiency of US and European R&I systems, particularly in relation to technological complexity. The visualisation employs a colour gradient where blue represents higher complexity technologies and red indicates lower complexity ones. A diagonal "competitiveness boundary" line serves as a critical reference point for understanding the performance gap.

Complex technologies (depicted in darker blue) consistently position themselves further above the diagonal boundary line. This indicates that these sophisticated technological domains suffer the most severe efficiency penalties due to the fragmentation of the European system, as compared to the US.

Figure 8. Complexity and R&I efficiency in EU/US (patents)

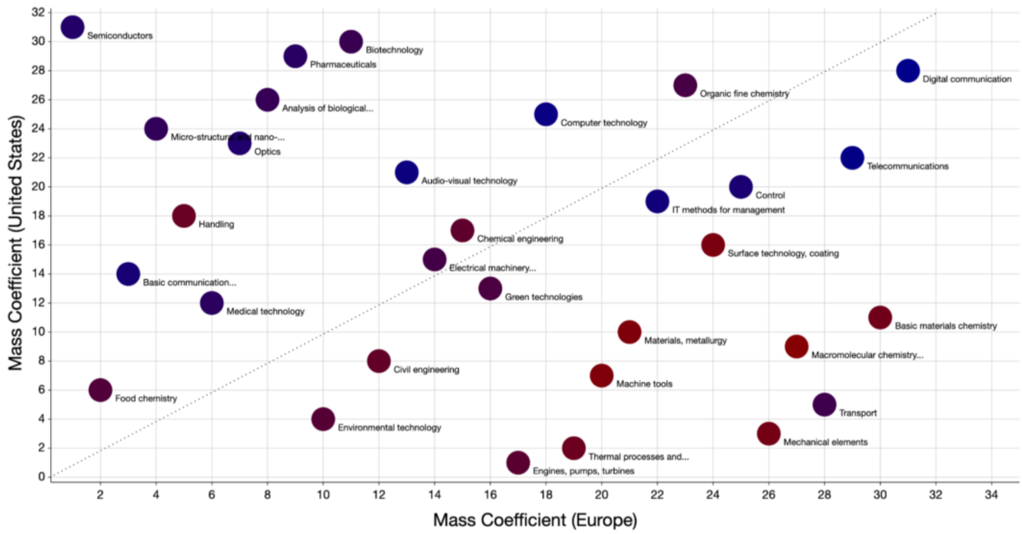
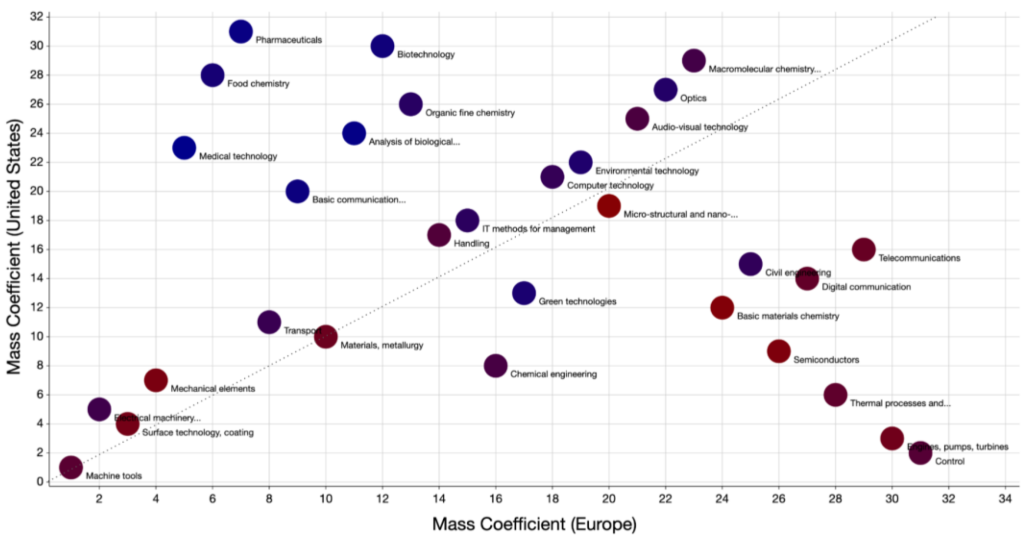


Figure 9 demonstrates that this pattern extends beyond patent activities into scientific publications; complex research areas show greater susceptibility to the efficiency losses associated with European fragmentation.

Figure 9. Complexity and R&I efficiency in EU/US (publications)



Overall, our results provide strong evidence in support of hypothesis 3, revealing a systematic relationship between technological complexity and the efficiency gap between the US and European systems. The pattern is not random, but rather shows a clear correlation: as technological complexity increases, the negative impact of European fragmentation becomes more pronounced. The parallel finding in both industrial innovation (patents) and academic research (publications) suggests that the fragmentation challenge in the European R&I system is structural and spans across different dimensions of R&I activities.

5. Robustness analysis

A number of supplementary analyses were conducted to examine the robustness of our findings to alternative geographical specifications. First, we redefine our geographical focus to Functional Urban Areas (FUA) and find consistent results for both patent and scientific collaborations (Figure A.1 to A.4).

Second, we examine the results at the NUTS3 level for Europe (Figure A.5 and A.8) and compare them to Core-Based Statistical Areas (CBSAs, Figure A.6 and A.9) and Territorial Level 3 (TL3 as defined by Fadic et al., 2019, Figure A.7 and A.10) regions in the US. CBSAs define urban-centred economic regions, while TL3 regions align with administrative boundaries, making them to a certain degree compatible with NUTS3 in terms of scale and economic relevance. However, unlike UA or FUA which are defined using a consistent methodology across the globe NUTS3, TL3 and CBSAs can slightly differ in coverage and definition.

Despite these differences, our results remain largely consistent across NUTS3, TL3, and CBSAs, with two notable variations. For patents, the same state coefficient is higher in CBSAs compared to NUTS3, suggesting a stronger within-border concentration of patenting activity in the US. For publications, Europe shows a higher mass coefficient compared to the US, suggesting more concentration of scientific output in European regions. These differences could potentially be attributed to the broader territorial coverage of NUTS3 in Europe, including all European regions and thus providing more opportunities for within-country collaborations. In contrast, CBSAs and TL3 regions mostly cover metropolitan and micropolitan areas thereby leading to a more concentrated collaboration pattern and potentially fewer opportunities for collaboration outside metropolitan areas.

6. Conclusion and policy implications

This paper provides compelling evidence that the fragmentation of the European R&I system represents a significant competitive disadvantage, particularly in complex technological domains that are crucial for future economic growth. Through our novel complexity-based approach and comprehensive analysis of patent and publication data, we have established three critical findings.

First, we conclusively demonstrate that the European R&I system exhibits substantially higher fragmentation compared to the US, with European hubs showing notably weaker interconnectivity than their US counterparts. This fragmentation is particularly pronounced in patent activities and remains evident, though less severe, in scientific publications. The key here is that the difference between the EU and the US stems from the EU's unrealized potential. Second, our analysis reveals that R&I systems for complex technologies inherently require greater integration and collaboration to function effectively. The strong correlation between technological complexity and R&I efficiency underscores that fragmentation is particularly costly in advanced technological domains. Third, and most crucially, we find that the Europe-US efficiency gap widens significantly for complex technologies, precisely where integration matters most. This "complexity penalty" of European fragmentation poses a serious threat to Europe's future competitiveness in strategic sectors like AI, biotechnology, and quantum computing.

These findings carry important policy implications. First, the framework proposed offers a new indicator able to capture the degree of fragmentation within R&I systems, providing policymakers with a new empirical tool to monitor changes in system cohesion over time, identify regions with high untapped collaboration potential, and devise strategies to foster more cohesive research and innovation across regions. The RISG can also be used to investigate pathways to a more inclusive development, by ensuring that smaller or peripheral regions are not being systematically excluded from the network, thereby enabling a fairer distribution of resources and collaboration opportunities based on each region's size, capabilities and overall connectivity. Additionally, the framework can be used to monitor the progress of policy initiatives, such as the ERA, over time, allowing for the assessment of their effectiveness in enhancing integration and collaboration across regions.

Second, our results call for decisive policy actions in shaping the next generation of European R&I programmes. This is particularly relevant for the reflections on the design of the next European Framework Programme for R&I. Our analysis demonstrates how strategic hub integration remains key to strengthen competitiveness, especially in strategic and more complex technologies. In this regard, R&I resources could be explicitly weighted toward projects that bridge multiple hubs in complex technological domains. Higher funding rates could be foreseen for multi-hub collaborative projects in complex technologies and dedicated budget lines could be designed for cross-border infrastructure sharing in advanced technological domains. This could include initiatives intended to create Complex Technology Integration Networks that provide sustained funding for multi-hub collaborations.

Beyond the Framework Programme for R&I, deeper structural changes are needed to reduce fragmentation through the development of pan-European research institutions with multiple hub locations, harmonisation of intellectual property rights and technology transfer procedures, as well as the creation of a true ERA with seamless mobility and resource sharing. As an example, establishing Hub Mobility Programs has the potential to facilitate researchers and innovators movement between major centres.

Complementary policies would also need to address broader market fragmentation through, for example, the acceleration of the Digital Single Market implementation, harmonisation of standards and regulations in emerging technology sectors, and the development of integrated European markets for research-intensive products and services.

The evidence suggests these reforms are not merely desirable, but essential for maintaining European competitiveness in an increasingly complex technological landscape. The cost of continued fragmentation, particularly in complex technologies, is simply too high to ignore. As the EU moves toward its next Framework Programme for R&I and beyond, policies need to prioritise the integration of European R&I hubs and the creation of a truly unified ERA capable of competing effectively in the most advanced technological domains.

7. References

Abbasiharofteh, M., Broekel, T., & Mewes, L. (2024). The role of geographic distance and technological complexity in U.S. interregional co-patenting over almost two centuries. *Environment and Planning A: Economy and Space*, 0(0).

Autant-Bernard, C., Billand, P., Frachisse, D., & Massard, N. (2007). Social distance versus spatial distance in R&D cooperation: Empirical evidence from European collaboration choices in micro and nanotechnologies. *Papers in Regional Science*, 86(3), 495-520.

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

Banchoff, T. (2002). Institutions, inertia and European Union research policy. *JCMS: Journal of Common Market Studies*, 40(1), 1-21.

Bathelt, H., Malmberg, A., & Clusters, P. (2004). Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography*, *28*(1), 31-56.

Boschma, R., (2005). Proximity and innovation: a critical assessment. *Reg. Stud*., 39 (1), 61–74

Bristow, G. (2005). Everyone's a ‘winner’: problematising the discourse of regional competitiveness. *Journal of economic geography*, 5(3), 285-304.

Cantwell, J., & Salmon, J., (2018). *Increasing knowledge complexity and informal networks in the information age*. In: Contractor, F.J., Reuer, J. (Eds.), Frontiers of Strategic Alliance Research: Negotiating, Structuring and Governing Partnerships. Cambridge University Press, Cambridge, NY. In press.

Chessa, A., Morescalchi, A., Pammolli, F., Penner, O., Petersen, A. M., & Riccaboni, M. (2013). Is Europe evolving toward an integrated research area?. *Science*, 339(6120), 650-651.

Crescenzi, R., Rodríguez-Pose, A., & Storper, M., (2007). On the geographical determinants of innovation in Europe and the United States. *J. Econ. Geogr.*, 7 (6), 673–709.

Draghi (2024). The future of European competitiveness – A competitiveness strategy for Europe.

Esposito, C. R. (2023). The geography of breakthrough invention in the United States over the 20th century. *Research Policy*, 52(7), 104810.

European Commission (2024). SRIP 2024 - Science, research and innovation performance of the EU – A competitive Europe for a sustainable future, Publications Office of the European Union, Luxembourg.

Fadic, M., Garcilazo, JE., Monroy, A., & Veneri, P. (2019). *Classifying small (TL3) regions based on metropolitan population, low density and remoteness*. OECD Regional Development Working Papers, No. 2019/06, OECD Publishing, Paris.

Fally, T., (2015). *Structural gravity and fixed effects*. NBER Working paper series, National bureau of economic research.

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

Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology*, 91(3), 481–510.

Hidalgo, C. A., & Hausmann, R., (2009). The building blocks of economic complexity. *Proc. Natl. Acad. Sci. U.S.A*., 106(12), 10570-10575.

Hoekman, J., Frenken, K., & van Oort, F., (2009). The geography of collaborative knowledge production in Europe. *The Annals of Regional Science*, 43(3), 721–738.

Jones, B., (2009). The burden of knowledge and the ‘death of the renaissance man’: Is innovation getting harder? *Review of Economic Studies,* 7(1), 283–317.

Lata, R., von Proff, S., & Brenner, T. (2017). The influence of distance types on co-patenting and co-publishing in the USA and Europe over time. *The Annals of Regional Science,* 31, 1069.

Maggioni, M. A., & Uberti, T. E. (2005). *International networks of knowledge flows: an econometric analysis* (No. 0519). Papers on Economics and Evolution.

Marshall, A. (1920). *Principle of Economics*. London: Macmillan

Maskell, P., & Malmberg, A. (1999). Localised learning and industrial competitiveness. *Cambridge Journal of Economics*, 23(2), 167-185.

Morescalchi, A., Pammolli, F., Penner, O., Petersen, A. M., & Riccaboni, M. (2015). The evolution of networks of innovators within and across borders: Evidence from patent data. *Research Policy*, 44(3), 651-668.

Morgan, K. (2007). The learning region: institutions, innovation and regional renewal. *Regional studies*, 41(1), 147-159.

Montobbio, F., & Sterzi,V. (2013). The globalization of technology in emerging markets: a gravity model on the determinants of international patent collaborations. *World Development*, 44, 281-299.

Okubo, Y., & Zitt, M. (2004). Searching for research integration across Europe: a closer look at international and inter-regional collaboration in France. *Science and Public Policy*, 31(3), 213-226.

Porter, M. E. (1990). The competitive advantage of notions. Harvard business review, 68(2), 73-93.

Rodriguez-Pose, A., & Crescenzi, R., (2008). Research and development, spillovers, innovation systems, and the genesis of regional growth in Europe. *Reg. Stud.* 42 (1), 51–67.

Shapira, P., Youtie, Y., Frietsch, R. (Ed.), & Schüller, M. (Ed.) (2010). *The Innovation System and Innovation Policy in the United States*. In Competing for Global Innovation Leadership: Innovation Systems and Policies in the USA, EU and Asia (pp. 5-20). Fraunhofer IRB.

Scherngell, T., & Barber, MJ. (2011). Distinct spatial characteristics of industrial and public research collaborations: Evidence from the fifth EU Framework Programme. *The Annals of Regional Science*, 46(2), 247–266.

Santos Silva, J., & Tenreyro, S. (2006). The log of gravity. *Rev Econ Stat,* 88, 641–658.

Santos Silva, J., & Tenreyro, S. (2022). The log of gravity at 15. *Portuguese Economic Journal*, 21(3).

Schmoch, U. (2008). Concept of a Technology Classification for Country Comparison. Final Report to the World Intellectual Property Organization.

Singh, J., & Marx, M. (2013). Geographic constraints on knowledge spillovers: political borders vs. spatial proximity. *Manag. Sci.*, 59 (9), 2056–2078.

United Nations, (2018). *World Urbanization Prospects*. Department of Economic and Social Affairs, Population Division, United Nations.

Van der Wouden, F. (2020). A history of collaboration in US invention: changing patterns of co-invention, complexity and geography. *Ind. Corp. Chang*. 29 (3), pp. 559–619.

WIPO (2019). *Tied in: The Global Network of Local Innovation*. Economic Research Working Paper No. 58/2019

8. Appendix

Table A.1. Econometric results for patent collaborations in Europe (FUAs)

	(1)	(2)	(3)	(4)
$Mass_{i,j,t}$	1.625***	1.635***	1.339***	1.398***
	(0.051)	(0.069)	(0.044)	(0.052)
$Dist_{i,j}$		-0.778***		-0.633***
		(0.044)		(0.047)
$SameCountry_{i,j}$			2.873***	1.677***
			(0.051)	(0.089)
Time Fixed Effects	Yes	Yes	Yes	Yes
No. Observations	508057	501138	508057	501138
Wald Chi-sq	1028.353	1396.208	3230.360	3984.986

Note: The continuous explanatory variables are taken in natural logarithms; their coefficients can be interpreted as elasticities. *, **, *** indicate significance at 10%, 5% and 1 %. Cluster-robust standard errors are shown in parentheses.

Table A.2. Econometric results for patent collaborations in the US (FUAs)

	(1)	(2)	(3)	(4)
$Mass_{i,j,t}$	2.512***	2.581***	2.495***	2.534***
	(0.105)	(0.099)	(0.102)	(0.102)
$Dist_{i,j}$		-0.518***		-0.282***
		(0.061)		(0.084)
$SameState_{i,j}$			1.526***	0.948***
			(0.126)	(0.190)
Time Fixed Effects	Yes	Yes	Yes	Yes
No. Observations	159179	159169	159179	159169
Wald Chi-sq	575.790	684.883	649.313	743.219

Note: The continuous explanatory variables are taken in natural logarithms; their coefficients can be interpreted as elasticities. *, **, *** indicate significance at 10%, 5% and 1 %. Cluster-robust standard errors are shown in parentheses.

Table A.3. Econometric results for scientific collaborations in Europe (FUAs)

	(1)	(2)	(3)	(4)
$Mass_{i,j,t}$	1.627***	1.563***	1.575***	1.591***
	(0.049)	(0.051)	(0.043)	(0.045)
$Dist_{i,j}$		-0.371***		0.046
		(0.018)		(0.035)
$SameCountry_{i,j}$			1.618***	1.665***
			(0.054)	(0.099)
Time Fixed Effects	Yes	Yes	Yes	Yes
No. Observations	118737	116503	118737	116503
Wald Chi-sq	1125.745	1485.711	1805.861	2288.794

Note: The continuous explanatory variables are taken in natural logarithms; their coefficients can be interpreted as elasticities. *, **, *** indicate significance at 10%, 5% and 1 %. Cluster-robust standard errors are shown in parentheses.

Table A.4. Econometric results for scientific collaborations in the US (FUAs)

	(1)	(2)	(3)	(4)
$Mass_{i,j,t}$	2.550***	2.550***	2.541***	2.534***
	(0.073)	(0.073)	(0.072)	(0.075)
$Dist_{i,j}$		-0.014		0.087
		(0.043)		(0.061)
$SameState_{i,j}$			0.387***	0.578***
			(0.099)	(0.171)
Time Fixed Effects	Yes	Yes	Yes	Yes
No. Observations	62091	62081	62091	62081
Wald Chi-sq	1226.765	1355.541	1301.983	1541.102

Note: The continuous explanatory variables are taken in natural logarithms; their coefficients can be interpreted as elasticities. *, **, *** indicate significance at 10%, 5% and 1 %. Cluster-robust standard errors are shown in parentheses.

Table A.5. Econometric results for patent collaborations in Europe (NUTS 3)

	(1)	(2)	(3)	(4)
$Mass_{i,j,t}$	2.207***	1.579***	1.892***	1.564***
	(0.032)	(0.018)	(0.031)	(0.017)
$Dist_{i,j}$		-1.203***		-1.088***
		(0.011)		(0.013)
$SameCountry_{i,j}$			2.887***	0.864***
			(0.023)	(0.033)
Time Fixed Effects	Yes	Yes	Yes	Yes
No. Observations	4178632	4178632	4178632	4178632
Wald Chi-sq	4780.741	16196.037	17423.474	37390.026

Note: The continuous explanatory variables are taken in natural logarithms; their coefficients can be interpreted as elasticities. *, **, *** indicate significance at 10%, 5% and 1 %. Cluster-robust standard errors are shown in parentheses.

Table A.6. Econometric results for patent collaborations in the US (CBSAs)

	(1)	(2)	(3)	(4)
$Mass_{i,j,t}$	2.008***	2.286***	1.962***	2.171***
	(0.092)	(0.085)	(0.078)	(0.079)
$Dist_{i,j}$		-1.054***		-0.798***
		(0.035)		(0.045)
$SameState_{i,j}$			3.333***	1.139***
			(0.107)	(0.098)
Time Fixed Effects	Yes	Yes	Yes	Yes
No. Observations	2587712	2587712	2587712	2587712
Wald Chi-sq	471.730	947.982	973.717	1423.454

Note: The continuous explanatory variables are taken in natural logarithms; their coefficients can be interpreted as elasticities. *, **, *** indicate significance at 10%, 5% and 1 %. Cluster-robust standard errors are shown in parentheses.

Table A.7. Econometric results for patent collaborations in the US (TL3)

	(1)	(2)	(3)	(4)
$Mass_{i,j,t}$	2.482***	2.581***	2.448***	2.536***
	(0.095)	(0.079)	(0.083)	(0.077)
$Dist_{i,j}$		-0.556***		-0.428***
		(0.042)		(0.054)
$SameState_{i,j}$			1.749***	0.687***
			(0.131)	(0.172)
Time Fixed Effects	Yes	Yes	Yes	Yes
No. Observations	159179	159179	159179	159179
Wald Chi-sq	688.337	1136.817	949.162	1214.971

Note: The continuous explanatory variables are taken in natural logarithms; their coefficients can be interpreted as elasticities. *, **, *** indicate significance at 10%, 5% and 1 %. Cluster-robust standard errors are shown in parentheses.

Table A.8. Econometric results for scientific collaborations in Europe (NUTS 3)

	(1)	(2)	(3)	(4)
$Mass_{i,j,t}$	4.134***	3.990***	4.157***	4.143***
	(0.077)	(0.067)	(0.054)	(0.055)
$Dist_{i,j}$		-0.472***		-0.058**
		(0.018)		(0.025)
$SameState_{i,j}$			1.885***	1.789***
			(0.041)	(0.062)
Time Fixed Effects	Yes	Yes	Yes	Yes
No. Observations	363581	363581	363581	363581
Wald Chi-sq	2879.272	4211.195	5901.049	6099.576

Note: The continuous explanatory variables are taken in natural logarithms; their coefficients can be interpreted as elasticities. *, **, *** indicate significance at 10%, 5% and 1 %. Cluster-robust standard errors are shown in parentheses.

Table A.9. Econometric results for scientific collaborations in the US (CBSAs)

	(1)	(2)	(3)	(4)
$Mass_{i,j,t}$	3.180***	3.180***	3.177***	3.174***
	(0.073)	(0.073)	(0.073)	(0.074)
$Dist_{i,j}$		0.009		0.121***
		(0.018)		(0.025)
$SameState_{i,j}$			0.622***	0.921***
			(0.112)	(0.160)
Time Fixed Effects	Yes	Yes	Yes	Yes
No. Observations	258967	258967	258967	258967
Wald Chi-sq	1897.971	1952.939	1998.436	2153.396

Note: The continuous explanatory variables are taken in natural logarithms; their coefficients can be interpreted as elasticities. *, **, *** indicate significance at 10%, 5% and 1 %. Cluster-robust standard errors are shown in parentheses.

Table A.10. Econometric results for scientific collaborations in the US (TL3)

	(1)	(2)	(3)	(4)
$Mass_{i,j,t}$	3.628***	3.634***	3.628***	3.632***
	(0.097)	(0.093)	(0.098)	(0.096)
$Dist_{i,j}$		-0.107***		-0.064
		(0.032)		(0.042)
$SameState_{i,j}$			0.535***	0.385***
			(0.097)	(0.142)
Time Fixed Effects	Yes	Yes	Yes	Yes
No. Observations	62091	62091	62091	62091
Wald Chi-sq	1385.904	1555.938	1488.899	1687.391

Note: The continuous explanatory variables are taken in natural logarithms; their coefficients can be interpreted as elasticities. *, **, *** indicate significance at 10%, 5% and 1 %. Cluster-robust standard errors are shown in parentheses.

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct centres. You can find the address of the centre nearest you online (european-union.europa.eu/contact-eu/meet-us_en).

On the phone or in writing

Europe Direct is a service that answers your questions about the European Union.
You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696,
- via the following form: european-union.europa.eu/contact-eu/write-us_en.

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website (european-union.europa.eu).

EU publications

You can view or order EU publications at op.europa.eu/en/publications. Multiple copies of free publications can be obtained by contacting Europe Direct or your local documentation centre (european-union.europa.eu/contact-eu/meet-us_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex (eur-lex.europa.eu).

EU open data

The portal data.europa.eu provides access to open datasets from the EU institutions, bodies and agencies. These can be downloaded and reused for free, for both commercial and non-commercial purposes. The portal also provides access to a wealth of datasets from European countries.

This study presents a novel complexity-based framework to analyse fragmentation in the EU's R&I system, highlighting hub connectivity as a critical factor. Drawing on extensive patent and publication data (2000–2023), it finds that European hubs are significantly less interconnected than their US counterparts, particularly in complex technologies such as AI, biotech, and quantum computing. The research underscores not only a performance gap but also structural inefficiencies, calling for more targeted, cross-regional policy interventions to enhance Europe's innovation competitiveness

Studies and reports

Publications Office
of the European Union