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

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

Papers in Evolutionary Economic Geography

26.08

Utrecht University Human Geography and Planning

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

Pierre-Alexandre BALLAND[1], Valentina DI GIROLAMO[2], Florence BENOIT[2], Julien RAVET[2], Alexandr HOBZA[2]

[1] Centre for European Policy Studies & Harvard University Growth Lab

[2] European Commission*, DG Research and Innovation

Abstract

Fragmentation in the European Union's R&I system is increasingly acknowledged as a major hindrance to its performance. However, 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 technological domains. Our analysis identifies 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 and scientific fields. Third, we find that there is an efficiency gap between the US and Europe in all domains, but it is most pronounced in complex ones, resulting in 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 and scientific domains.

Keyword: Research Fragmentation, Economic Complexity, Inter-Regional Connectivity, Innovation Networks, EU Competitiveness.

* The views expressed in this paper are those of the authors involved and may, under no circumstances, be interpreted as stating an official position of the European Commission.

1. Introduction

We may be experiencing the most rapid and transformative technological revolution in human history. Advances in fields such as artificial intelligence (AI), biotechnology and advanced materials are reshaping production systems, global value chains and the overall geography of innovation. In this context, the performance of research and innovation (R&I) systems has become a central determinant of economic success and global influence. Indeed, the ability of economies to generate, absorb and recombine knowledge is increasingly shaping their capacity to sustain productivity growth and technological leadership.

Innovation increasingly emerges not from isolated actors but from complex, interconnected networks of firms, universities, research institutes, and public agencies. These networks drive the generation, recombination, and diffusion of knowledge, which are essential for developing novel technologies (Balland et al., 2018). The effectiveness of R&I systems therefore depends not only on the capabilities of individual organisations but also on the connectivity of knowledge actors across geographic and institutional boundaries. Spatial proximity, localised interactions, and the degree to which system components are connected strongly influence innovation performance. Fragmentation of the system, reflected in weak connections, limits knowledge flows and coordination, duplicates efforts and reduces the overall effectiveness of the system (Draghi, 2024; van der Pol & Frenken, 2025). In other words, the innovation potential of the system remains underutilised. . This can have important implications for technologically complex domains, where innovation requires the integration of multidisciplinary knowledge that is often distributed across multiple locations (Hidalgo & Hausmann, 2009; Balland & Rigby, 2017).

The issue of fragmentation is particularly salient in Europe. Despite considerable collective resources and talent, the European innovation landscape remains characterised by persistent institutional and regulatory barriers, an issue that has been addressed in various high-level policy reports (e.g., Draghi, 2024; Letta, 2024). Within Europe, each Member State largely pursues its own innovation priorities, shaped by local economic needs and political agendas. The absence of a more coordinated approach hampers the integration of European research hubs, reduces opportunities for collaborative projects and weakens the system's capacity to generate and recombine knowledge at scale. This, in turn, reduces Europe's ability to leverage network effects, scale strategically important technologies, and maintain competitiveness in a global innovation ecosystem increasingly dominated by winner-takes-all dynamics. By contrast, the US has benefited from a more integrated R&I system, with a unified national strategy and federal programmes that actively foster collaboration and knowledge sharing across States, facilitating more effective hub connectivity and network integration. This contrast makes the US a natural benchmark to investigate the effects of fragmentation on innovation performance.

Despite the increasing prominence of fragmentation in policy debates, academic literature has only partially addressed this issue. A large body of research has examined the geography of innovation networks and the determinants of collaborative knowledge production, highlighting the importance of spatial proximity, institutional similarity and network structures for knowledge diffusion and innovation performance (e.g. Boschma, 2005; Hoekman et al., 2010; Balland & Rigby, 2017). Related work has also explored the spatial concentration of complex technologies and the role of innovation hubs in global knowledge networks (Hidalgo & Hausmann, 2009; Balland et al., 2020). However, existing studies have largely focused on the determinants of collaboration or the spatial concentration of innovative activities, rather than directly assessing the degree of fragmentation of R&I systems as a systemic property. Moreover, relatively little attention has been devoted to how fragmentation interacts with technological complexity, despite growing evidence that complex technologies rely more heavily on geographically distributed and multidisciplinary knowledge bases.

This paper aims to address this gap by introducing a novel complexity-based framework to measure fragmentation in R&I systems. Our approach focuses on the connectivity between major innovation hubs, capturing the extent to which key centres of R&I collaborate and exchange knowledge. Using a comprehensive dataset combining patent applications from the OECD REGPAT database and scientific publications from OpenAlex between 2000 and 2023, we map collaboration networks across multiple spatial levels and technological fields. This multi-scalar approach allows us to capture the multifaceted nature of R&I networks.

Our analysis highlights three key findings. First, the European R&I system is significantly more fragmented compared to that of the US, with major European hubs showing notably weaker interconnectivity. Second, hub connectivity becomes increasingly important as technological complexity rises. Third, fragmentation in Europe is particularly pronounced in complex technological domains, suggesting that institutional and spatial barriers disproportionately affect the development of advanced technologies.

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

R&I system fragmentation, and introducing the hypotheses tested in this paper. Section 3 describes the data used 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 and discusses the policy implications of the results.

2. Theoretical framework

A core argument of evolutionary economics is that economic growth is driven by dynamic processes of technological change, learning, and knowledge accumulation, rather than by static equilibrium conditions. Building on these perspectives, the innovation systems literature conceptualises innovation as an interactive and systemic process that emerges from interactions among firms, research organisations, universities, and governments rather than from isolated actors (e.g., Lundvall, 2007; Edquist, 2010). Innovation can, thus, be understood as a networked process, where the positions of the different actors within networks influence the way knowledge spreads across the system (Fritsch & Kauffeld-Monz, 2010). Because knowledge is dispersed across firms and organisations, collaboration networks play a critical role, acting as channels for knowledge exchange and enabling different actors to access complementary expertise and technologies (Iino et al., 2021; Balland et al., 2018). The resulting creation and recombination of knowledge and capabilities are key drivers for technological change and productivity improvements, ultimately playing a central role in fostering economic growth.

In this regard, spatial proximity, localised interactions and, more broadly, the way different components of an R&I system connect and interact with each other determine its overall innovation performance. Weak connections can lead to fragmentation, limit knowledge flows and coordination, and reduce the effectiveness of the whole system (van der Pol & Frenken, 2025). These considerations are particularly relevant when assessing the case of the EU R&I system. The latter has long suffered from a significant degree of fragmentation, largely stemming from the institutional architecture of research governance in Europe, where competencies for R&I policy remain primarily at the national level, while coordination at the supranational level has historically been comparatively limited.

Within the EU, each Member State largely pursues its own innovation priorities, shaped by local economic needs and political agendas. As a result, funding for R&I is often dispersed and misaligned across Member States, while only about 10% of total R&I spending is directly managed at EU level through EU programmes (Draghi, 2024). The lack of coordination often translates into multiple entities independently pursuing similar projects, typically leading to small-scale efforts, and to missed opportunities for network effects. Fragmentation limits Europe's ability to concentrate resources on a cohesive set of key technologies essential for global competitiveness and prevents the creation of pan-European networks. This, in turn, limits cross-border collaboration and knowledge-sharing, needed for scaling critical and most advanced technologies, de facto hindering Europe's ability to compete in a global technological ecosystem, in many technological areas driven by a winner-takes-all model. In sharp contrast, the US benefits from a more integrated R&I system, with a large, unified market supported by federal programmes and a single national strategy, that actively foster collaboration and knowledge sharing across States, thereby ensuring a more coordinated allocation of resources and a clearer alignment between investment priorities.

Against this background and given its well-established position as a global technological leader, the US offers a natural benchmark against which to investigate the structure and integration of the European R&I system. In the following sections, we develop a set of hypotheses on the geography of the R&I collaborative networks in both Europe and the US, and on how differences in the degree of fragmentation can affect the development of complex technologies and scientific fields.

2.1 The geography of R&I collaborative networks in Europe and the 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 a 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 countries, which pursue their own R&I policies often driven by diverse national priorities. Thanks to the fact that R&I policies are a shared competence, the EU institutions attempt to support and coordinate national policies. Despite this the European science and innovation space remains highly heterogeneous. Such heterogeneity manifests itself in each Member State having its own research agenda, funding mechanisms, and policy frameworks and can lead to duplication of efforts and lack of necessary scale, particularly as regards complex technologies which require large, coordinated investments.

Scientific and innovation activities are territorially embedded processes that need specific structural and institutional conditions in order 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; von Proff & Brenner, 2014; 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 significant 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 centralised 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 implies thatthe European innovation system islargely characterisedby distinct national and regional institutions, with diverse priorities aimed 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 and involve 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 to 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 (Marshall, 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 20^{th} 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 depends more strongly on local interactions and intense collaborations than less complex one (e.g., Balland & Rigby 2017; Broeckle, 2019; 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 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 tends to cluster, accessing expertise that is 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 influence where R&I collaborations occur and how they structure geographically 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, Esposito & Van der Wouden, 2022). 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 test 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 postulate an additional hypothesis. If more complex technologies are disproportionately affected by R&I system fragmentation, then the innovative gap in Europe should be concentrated in these technologies. Simpler technologies, which rely less on unique combinations of knowledge, suffer to a lesser extent from barriers to spatial collaborations.

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. 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 and methodology

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 European data cover the 27 Member States of the EU, the UK and the EFTA countries. This broad scope is driven by 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, and excluding them would overlook key nodes of the European collaborative network, potentially distorting the overall picture of European R&I activity. Third, ensuring a broader understanding of European R&I collaborative activity when benchmarking it against the US remains key also from a policy perspective, allowing the analysis to deliver more comprehensive insights to inform strategic policy decisions.

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

3.1.2 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 categorised 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 provide 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 co-authorship on the same publication. Publications were categorised into the WIPO 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 categorisation, where the LLM integrated the embedding similarity and relatedness density metrics to make informed classification decisions. Finally, to minimise false positives, we manually defined a threshold for inclusion, ensuring that only robust and meaningful topiccategory 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 Introducing the Research & Innovation System Gap indicator

To quantify the fragmentation and efficiency of the R&I systems, the paper introduces a new indicator, namely the Research & Innovation System Gap (RISG). It measures the correlation between the 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 one link within Santa Clara, CA. But we do not double count the two links between London and Santa Clara, CA.

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.

Figure 1. Counting links

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 weight 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 deviate from the expected distribution, 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 less connected than 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.

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 R&I systems.

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 distance and differences in preferences and specialisations (Burger et al., 2007). Additionally, the PPML estimator offers a more tractable empirical strategy for incorporating fixed effects and handling multilateral resistance terms[1], 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:

$$(1)\ Geo\ Links_{i,j,c,t} = \alpha + \beta_1 Mass_{i,j,t} + \beta_2 Dist_{i,j} + \beta_4 Same\ Country_{i,j} + \omega_t + \varepsilon_{i,j,t}$$

where $Geo\ Links_{i,j,c,t}$ is a count variable indicating the number of (patent or publication) collaboration links between urban area i and urban area j, for a given tech/scientific domain c,

[1] 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.

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.

$Mass_{i,j,c,t}$ 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):

$$(2)\ Geo\ Links_{i,j,c,t} = \alpha + \beta_1 Mass_{i,j,c,t} + \beta_2 Dist_{i,j} + \beta_3\, Tech\ Proximity_{i,j,c,t} + \\ + \beta_4 Same\ Country_{i,j} + \omega_c + \omega_t + \varepsilon_{i,j,t}$$

The additional control $Tech\ Proximity_{i,j,c,t}$ captures the technological proximity between i *and* j. Technological proximity reflects the similarity in knowledge capabilities and absorptive capacity[2] 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). $Same\ Country_{i,j}$ 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.

[2] 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.

All continuous variables were logarithmically transformed. This 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. The descriptives of the sample by WIPO categories can be found in the appendix.

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 Europe for patents and in the US for publications.

Variable	Description	Patent data				Publication data			
		Europe		US		Europe		US	
		Mean	*SD.*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Geo Links	*Number of collaborations between two spatial areas, for a given tech/scientific domain.*	1.57	13.38	8.70	49.62	215.21	926.28	160.61	767.27
Mass	*Number of collaborations two spatial areas are engaged in.*	837.59	929.13	3,086.75	3,369.32	100,120.00	153,838.50	90,250.08	119,040.60
Distance	*Geographical distance (km) between two geographical areas.*	1,101.58	622.89	1,861.51	1,244.81	1,851.51	1,238.02	1,107.47	626.30
Same country/ Same State	*Dummy variable taking value 1 if two geographical areas are located within the same country or State in Europe and the US, respectively.*	0.10	0.31	0.05	0.21	0.04	0.21	0.10	0.30

Table 1. Descriptive Statistics

Table 2 indicates a positive correlation between mass and geographical links, which is stronger in the US for patents and publications than in Europe. Distance negatively correlates with geographical 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.

Patent data

Variable	Europe				US			
Geo Links	1.0000				1.000			
Mass	0.2618	1.000			0.3653	1.000		
Distance	-0.1408	-0.2134	1.000		-0.0558	-0.0719	1.000	
Same country/Same State	0.2424	0.0168	-0.4065	1.000	0.1306	0.0316	-0.2763	1.000

Publication data

Variable	Europe				US			
Geo Links	1.0000				1.000			
Mass	0.4710	1.000			0.6223	1.000		
Distance	-0.1214	-0.0606	1.000		-0.0263	0.0095	1.000	
Same country/Same State	0.2210	-0.0116	-0.4049	1.000	0.0109	-0.0212	-0.2716	1.000

Table 2. Correlations

4. Results and discussion

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

Patent data

The US demonstrates a more cohesive R&I system compared to Europe, consistent with Hypothesis 1. The US system exhibits a strong RISG[3] 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

[3] 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.

through network analysis[4], showing connections that span across the entire country without significant border-related clustering (Figure 2[5]).

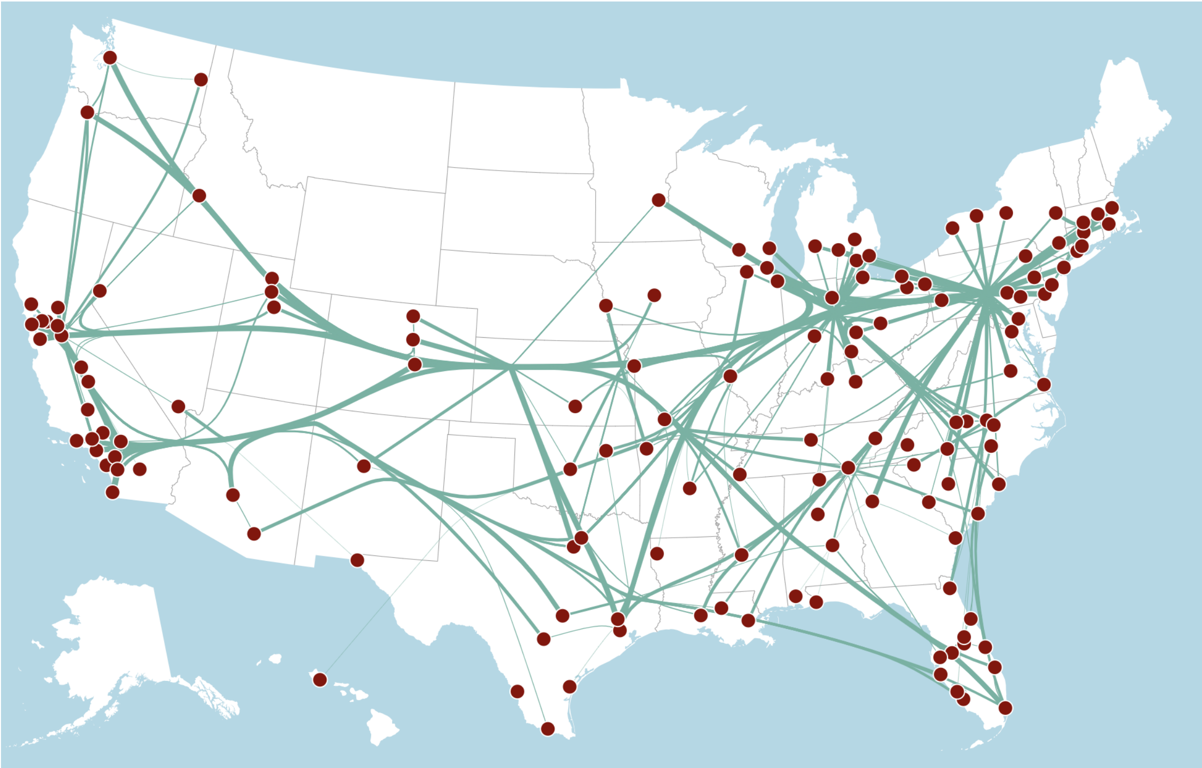

Figure 2. R&I systems in the US, patent data

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 along 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, are likely to play a more significant role in shaping European innovation partnerships.

[4] 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.

[5] Tickers lines indicate larger amount of links.

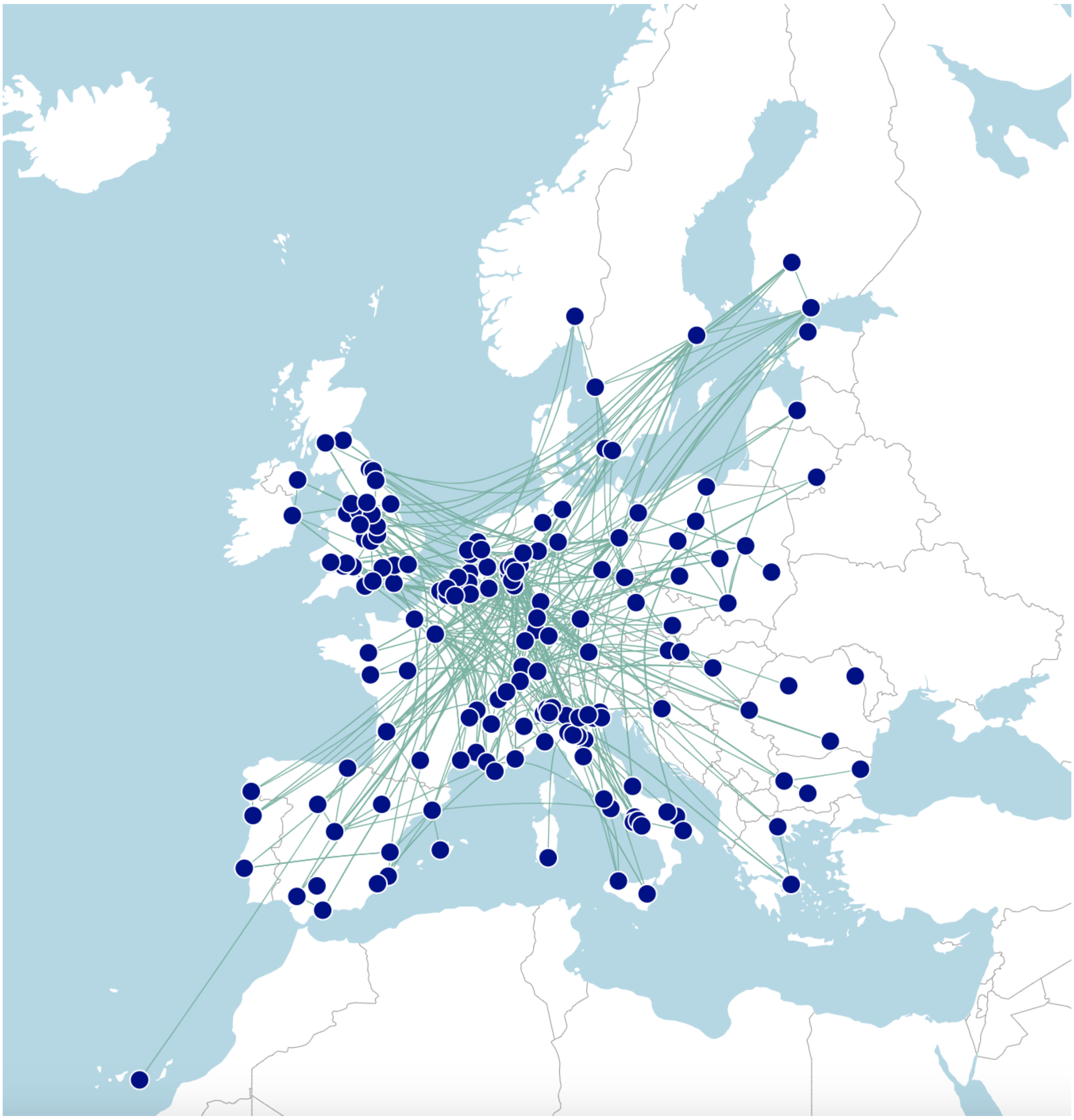

Figure 3. R&I systems in Europe, patent data

Scientific publications

The scientific publication landscape reveals a somewhat different picture, while still showing a 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).

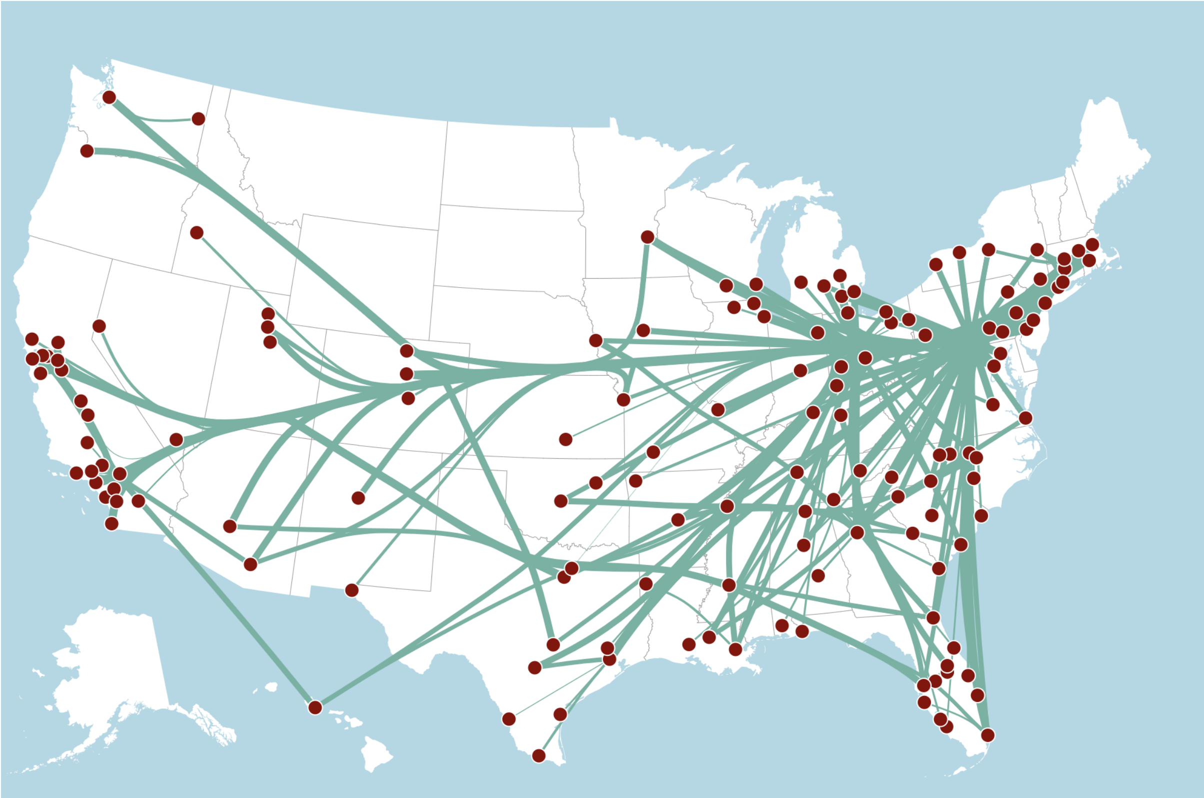

Figure 4. R&I systems in the US, publication data

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).

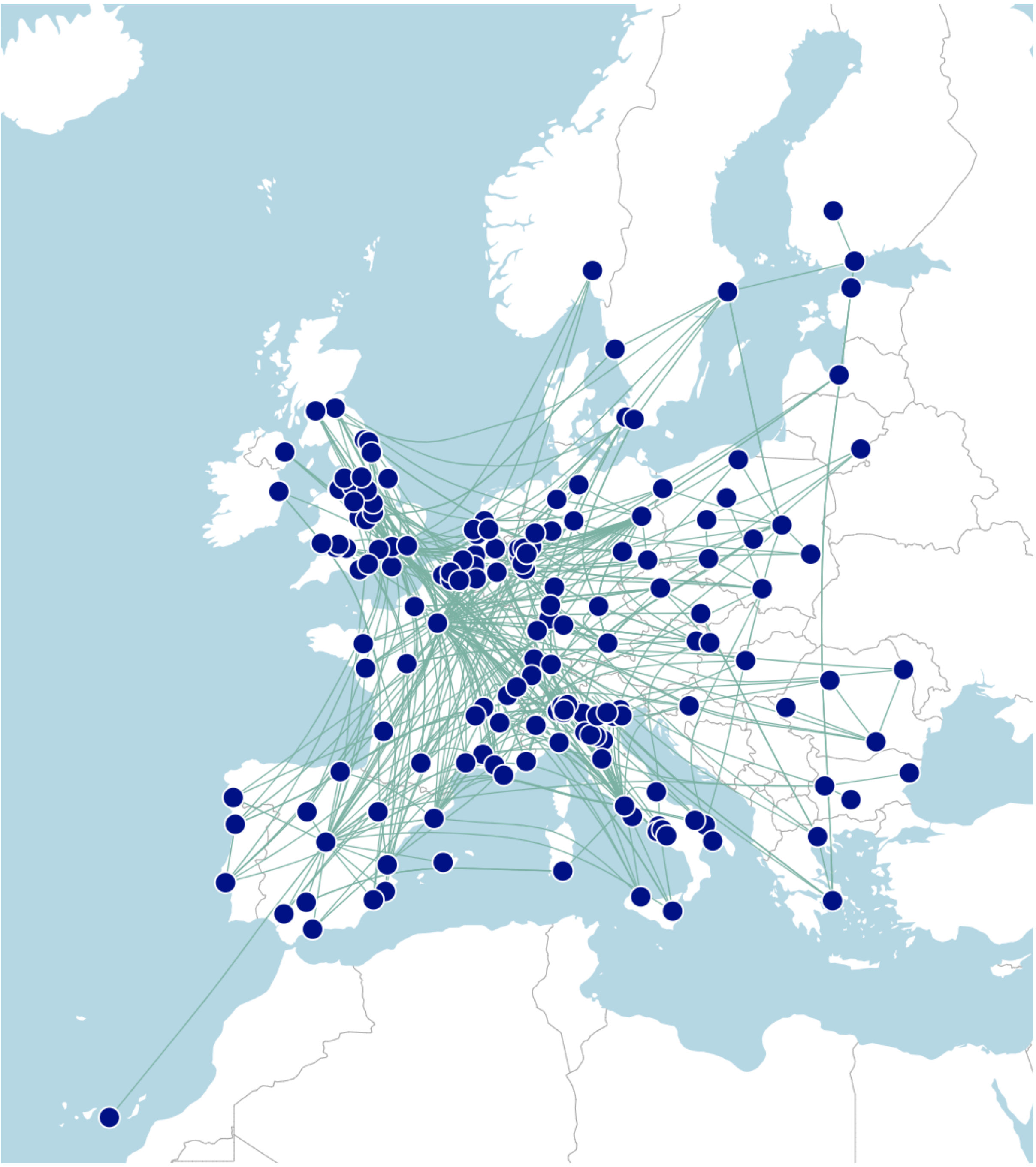

Figure 5. R&I systems in the Europe, publication data

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. Tables 3 to 6 report the results from estimating specification (1) for patent and publication data, respectively.

	(1)	(2)	(3)	(4)
Mass ij	2.405***	1.928***	1.868***	1.799***
	(0.117)	(0.061)	(0.065)	(0.049)
Distance ij		-0.901***		-0.554***
		(0.018)		(0.026)
Same Country ij			2.981***	1.852***
			(0.062)	(0.083)
Time Fixed Effects	Yes	Yes	Yes	Yes
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 3. Econometric results for patent collaborations in Europe

	(1)	(2)	(3)	(4)
Mass ij	1.989***	2.009***	1.904***	1.997***
	(0.078)	(0.057)	(0.061)	(0.054)
Distance ij		-0.494***		-0.472***
		(0.031)		(0.026)
Same State ij			1.303***	0.132
			(0.134)	(0.083)
Time Fixed Effects	Yes	Yes	Yes	Yes
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.*

Table 4. Econometric results for patent collaborations in the US

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 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 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 (Europe) or same State (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 State boundaries, reflecting a more integrated innovation system.

	(1)	(2)	(3)	(4)
Mass ij	2.313***	2.294***	2.272***	2.281***
	(0.052)	(0.051)	(0.034)	(0.036)
Distance ij		-0.519***		-0.131***
		(0.014)		(0.019)
Same Country ij			1.884***	1.671***
			(0.037)	(0.055)
Time Fixed Effects	Yes	Yes	Yes	Yes
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 5. Econometric results for scientific collaborations in Europe

	(1)	(2)	(3)	(4)
Mass ij	2.911***	2.891***	2.925***	2.898***
	(0.052)	(0.049)	(0.052)	(0.049)
Distance ij		-0.165***		-0.153***
		(0.020)		(0.023)
Same State ij			0.504***	0.155
			(0.083)	(0.101)
Time Fixed Effects	Yes	Yes	Yes	Yes
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.*

Table 6. Econometric results for scientific collaborations in the US

In terms of scientific collaborations, the insights from the RISG analysis are corroborated by the econometric results (Tables 5 and 6). Although the US continues to report higher coefficients for *mass*, 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 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[6] and the mass coefficients derived from separate regression analyses for each technology category.[7] The mass coefficient (as 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.

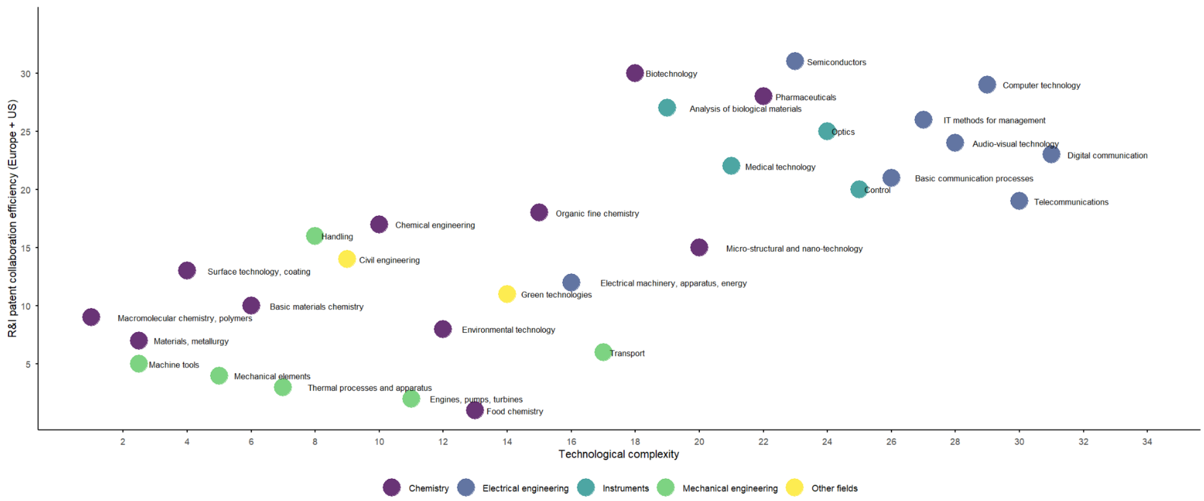

Figure 6. Complexity and R&I efficiency, patent data

[6] 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.

[7] 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).

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 consistent with Hypothesis 2.

These results highlight the role of pan-European collaborative efforts, especially in advanced technological fields, in supporting 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 scope for targeted strategies to enhance R&I efficiency in sectors that currently underperform relative to their complexity level.

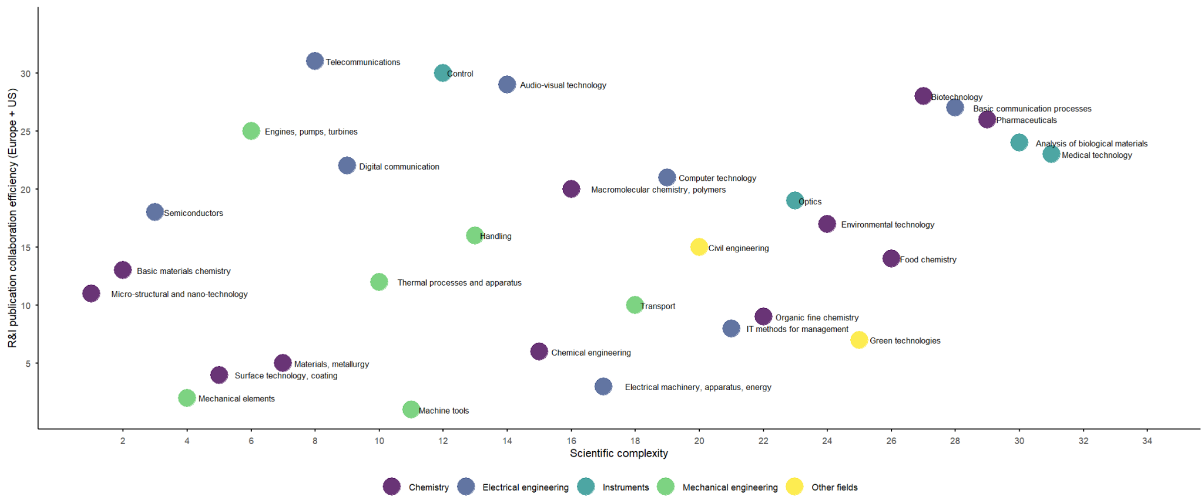

Figure 7. Complexity and R&I efficiency, publication data

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.[8] 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 darker shades indicated higher complexity technologies and scientific fields. A diagonal "competitiveness boundary" line serves as a reference point for understanding the performance gap.

[8] As before, the coefficient estimates are obtained by estimating specification (2) as outlined in Section 5.

Complex technologies consistently lie 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.

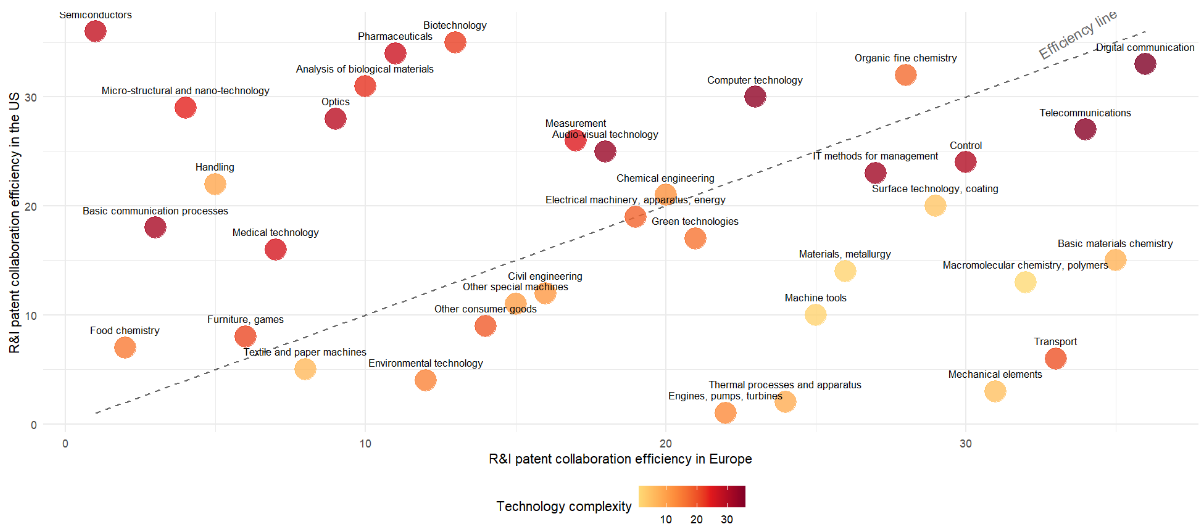

Figure 8. The relationship between complexity and R&I collaboration efficiency in Europe and the US, patent data

Figure 10 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.

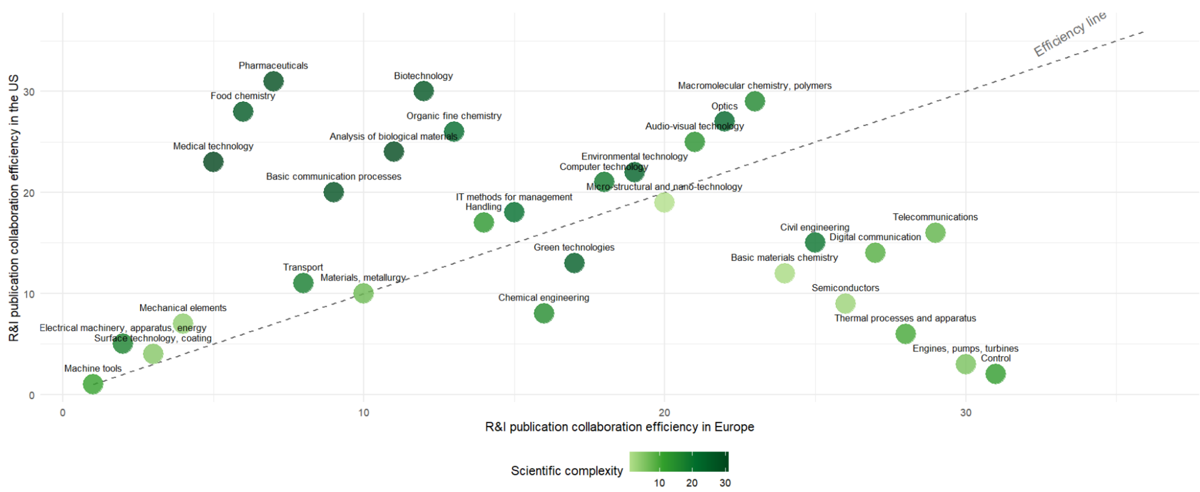

Figure 9. The relationship between complexity and R&I collaboration efficiency in Europe and the US, publication data

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.

Second, we examine the results at the NUTS3 level for Europe and compare them to Core-Based Statistical Areas (CBSAs) and Territorial Level 3 (TL3 as defined by Fadic et al., 2019) regions in the US. CBSAs define urban-centered 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 a higher 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

This paper investigates the geography of collaborative R&I networks in Europe and the US with a focus on their respective degree of fragmentation. We build on the innovation systems literature and on insights from evolutionary economics, introducing a novel indicator (the RISG) to measure the degree of fragmentation within an R&I system. We also investigate how fragmentation interacts with technological complexity, showing that the lack of integration represents an important competitive disadvantage in more complex technological domains. Specifically, through our novel complexity-based approach and comprehensive analysis of patent and publication data, we have established three critical findings.

First, we 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. 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, we find that the Europe-US efficiency gap widens significantly for complex technologies, precisely where integration matters most. This patterns suggests a "complexity penalty" associated with European fragmentation, particularly in strategic sectors.

Our findings carry important policy implications. First, our framework 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 integration over time. Additionally, the framework can be used to monitor the progress of policy initiatives over time, allowing for the assessment of their effectiveness in enhancing integration and collaboration across regions. Second, our analysis demonstrates how strategic hub integration remains key to strengthen competitiveness, especially in more complex fields. In this regard, R&I resources could be explicitly weighted toward projects that bridge multiple hubs in complex technological domains. Higher funding rates may be considered for multi-hub collaborative projects in complex technologies alongside support for cross-border infrastructure sharing in advanced technological domains.

Despite the robustness of our findings, some caveats deserve attention. First, part of the empirical analysis relies on patent data to capture collaborative innovation. While patents provide a well-established proxy to measure innovation activities and technological collaboration, they do not fully capture all forms of innovation. While we complement patent data with publication data, some dimensions of collaboration and knowledge exchange may remain only partially observed. Second, the parametric validation of the RISG findings relies on a relatively limited set of control variables. This reflects both a deliberate focus on structural differences in the organisation of R&I systems and the high spatial granularity of our network analysis, which is conducted at the urban-area level and, thus, makes it difficult to construct a broader set of comparable institutional, financial, and policy indicators across both the European and US systems. Third, cross-regional comparisons inevitably depend on the definition of the geographical units adopted in the analysis. In order to ensure that our findings are not driven by the specific geographical definition adopted in the baseline specification, we conducted a series of robustness checks using alternative spatial classifications, i.e. FUA, NUTS3, CBSA, and TL3 regions. While our results remain broadly consistent across these alternative specifications, we acknowledge that these territorial definitions are not perfectly harmonised across the European and the US systems, which may affect observed collaboration patterns.

At last, it is important to highlight that the framework proposed in this paper and, in particular, the RISG indicator also present some limitations that should be taken into account when interpreting the results. A first limitation relates to the scale-invariant nature of the RISG indicator. The latter is constructed as a correlation coefficient between observed and expected collaboration links. As such, it captures the relative structure of the network rather than its absolute scale, remaining insensitive to absolute network size and density. As a result, the RISG applied to two R&I systems that vastly differ in terms of number of links, hubs or researchers can in principle yield identical correlation scores, making comparisons more challenging. In our analysis, we partially compensate for this weakness by considering spatial units of analysis that are comparable between Europe and the US, thereby making the two systems reasonably similar in scale and, thus, suitable for comparison. Nevertheless, researchers seeking to extend this approach to other contexts should take this property into account. A second limitation concerns the specification of the null model used to derive expected collaboration links, which conflates structural constraints with policy-relevant determinants of fragmentation. Specifically, in our framework, expected links are determined solely by the relative size of the hubs, while other factors that also shape collaboration patterns (e.g., geographical distance, language, institutional compatibility, or regulatory environments) are not factored in. On the one hand, such a parsimonious specification makes the RISG an agile indicator able to measure fragmentation based on a simple size-based benchmark. However, on the other hand, it also implies that it cannot distinguish between different sources of fragmentation. This is a relevant point, as some of these sources, such as national borders, different regulatory frameworks, or heterogeneous intellectual property (IP) regimes, are the barriers that policy interventions typically aim to address in order to strengthen integration within an R&I system. The RISG can only capture fragmentation in a broad structural sense as a deviation from an optimal size-proportional network, but cannot identify what part of this deviation corresponds to policy-relevant barriers. Furthermore, our indicator summarises the structure of the entire collaboration network into a single scalar. While this aggregation has the advantage of facilitating comparisons across systems, it also inevitably masks the specific geography of fragmentation, including which hub pairs are more under-connected, where the largest deviations from expected collaboration occur, and whether integration is improving in the areas that matter most for strategic technological development. As a result, future avenues of research could focus on developing a more granular decomposition to better identify specific intervention points. Finally, while our results provide strong evidence on the structural differences in the degree of fragmentation between the European and the US R&I collaborative systems, this paper does not attempt to directly quantify the economic cost of fragmentation. Building on the existing academic literature on innovation systems, our analysis postulates that fragmentation represents a constraint to the efficiency of collaborative networks and, therefore, on the development of complex technologies. However, the empirical framework developed in this paper primarily focuses on measuring the degree of fragmentation within R&I collaborative systems and not on estimating its direct impact on indicators such as productivity, innovation performance, or economic growth. Future research could build on the RISG framework to more explicitly investigate these aspects, and explore how differences in network integration translate into measurable economic outcomes across regions, sectors, and technological domains.

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.

van der Pol, J., & Frenken, K. (2025). Fragmentation of national research systems: the case of the Netherlands. Scientometrics, 130(6), 3213-3227.

Balland, P.A., Jara-Figueroa, C., Petralia, S., Steijn, M., Rigby, D., and Hidalgo, C. (2020) Complex Economic Activities Concentrate in Large Cities, Nature Human Behavior, 4: 248–254

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

Bathelt, H., Malmberg, A., & Maskell, 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.

Edquist, C. (2010). Systems of innovation perspectives and challenges. African journal of science, Technology, Innovation and Development, 2(3), 14-45.

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." Journal of International Economics, 97(1), 76–85.

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.

Fritsch, M., & Kauffeld-Monz, M. (2010). The impact of network structure on knowledge transfer: an application of social network analysis in the context of regional innovation networks. The Annals of Regional Science, 44(1), 21-38.

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), pp. 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, Benjamin F. (2009). “The burden of knowledge and the ‘death of the renaissance man’: Is innovation getting harder?” Review of Economic Studies 7, 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.

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(S1), S147-S159.

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.

Iino, T., Inoue, H., Saito, Y. U., & Todo, Y. (2021). How does the global network of research collaboration affect the quality of innovation?. The Japanese Economic Review, 72(1), 5-48.

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.

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, pp. 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), pp. 2056–2078.

United Nations, Department of Economic and Social Affairs, Population Division (2019). World Urbanization Prospects. New York: United Nations.

Esposito, C., & Van Der Wouden, F. (2022). Learning Amid Technological Disruption: Evidencefrom US Inventors, 1836-1999. In Academy of Management Proceedings (Vol. 2022, No. 1, p. 17898). Briarcliff Manor, NY 10510: Academy of Management.

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.

von proff, S., & Brenner, T. (2014). The dynamics of inter-regional collaboration: an analysis of co-patenting. The Annals of Regional Science 52, pp. 41-64.

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

APPENDIX

	(1)	(2)	(3)	(4)
Mass ij	1.625***	1.635***	1.339***	1.398***
	(0.051)	(0.069)	(0.044)	(0.052)
Distance ij		-0.778***		-0.633***
		(0.044)		(0.047)
Same Country ij			2.873***	1.677***
			(0.051)	(0.089)
Time Fixed Effects	Yes	Yes	Yes	Yes
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.1. Econometric results for patent collaborations in Europe (FUAs)

	(1)	(2)	(3)	(4)
Mass ij	2.512***	2.581***	2.495***	2.534***
	(0.105)	(0.099)	(0.102)	(0.102)
Distance ij		-0.518***		-0.282***
		(0.061)		(0.084)
Same State ij			1.526***	0.948***
			(0.126)	(0.190)
Time Fixed Effects	Yes	Yes	Yes	Yes
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.2. Econometric results for patent collaborations in the US (FUAs)

	(1)	(2)	(3)	(4)
Mass ij	1.627***	1.563***	1.575***	1.591***
	(0.049)	(0.051)	(0.043)	(0.045)
Distance ij		-0.371***		0.046

		(0.018)		(0.035)
Same Country ij			1.618***	1.665***
			(0.054)	(0.099)
Time Fixed Effects	Yes	Yes	Yes	Yes
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.3. Econometric results for scientific collaborations in Europe (FUAs)

	(1)	(2)	(3)	(4)
Mass ij	2.550***	2.550***	2.541***	2.534***
	(0.073)	(0.073)	(0.072)	(0.075)
Distance ij		-0.014		0.087
		(0.043)		(0.061)
Same State ij			0.387***	0.578***
			(0.099)	(0.171)
Time Fixed Effects	Yes	Yes	Yes	Yes
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.4. Econometric results for scientific collaborations in the US (FUAs)

	(1)	(2)	(3)	(4)
Mass ij	2.207***	1.579***	1.892***	1.564***
	(0.032)	(0.018)	(0.031)	(0.017)
Distance ij		-1.203***		-1.088***
		(0.011)		(0.013)
Same Country ij			2.887***	0.864***
			(0.023)	(0.033)
Time Fixed Effects	Yes	Yes	Yes	Yes
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.5. Econometric results for patent collaborations in Europe (NUTS 3)

	(1)	(2)	(3)	(4)
Mass ij	2.008***	2.286***	1.962***	2.171***
	(0.092)	(0.085)	(0.078)	(0.079)
Distance ij		-1.054***		-0.798***
		(0.035)		(0.045)
Same State ij			3.333***	1.139***
			(0.107)	(0.098)
Time Fixed Effects	Yes	Yes	Yes	Yes
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.6. Econometric results for patent collaborations in the US (CBSAs)

	(1)	(2)	(3)	(4)
Mass ij	2.482***	2.581***	2.448***	2.536***
	(0.095)	(0.079)	(0.083)	(0.077)
Distance ij		-0.556***		-0.428***
		(0.042)		(0.054)
Same State ij			1.749***	0.687***
			(0.131)	(0.172)
Time Fixed Effects	Yes	Yes	Yes	Yes
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.7. Econometric results for patent collaborations in the US (TL3)

	(1)	(2)	(3)	(4)

Mass ij	4.134***	3.990***	4.157***	4.143***
	(0.077)	(0.067)	(0.054)	(0.055)
Distance ij		-0.472***		-0.058**
		(0.018)		(0.025)
Same Country ij			1.885***	1.789***
			(0.041)	(0.062)
Time Fixed Effects	Yes	Yes	Yes	Yes
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.8. Econometric results for scientific collaborations in Europe (NUTS 3)

	(1)	(2)	(3)	(4)
Mass ij	3.180***	3.180***	3.177***	3.174***
	(0.073)	(0.073)	(0.073)	(0.074)
Distance ij		0.009		0.121***
		(0.037)		(0.045)
Same State ij			0.622***	0.921***
			(0.112)	(0.160)
Time Fixed Effects	Yes	Yes	Yes	Yes
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.9. Econometric results for scientific collaborations in the US (CBSAs)

	(1)	(2)	(3)	(4)
Mass ij	3.628***	3.634***	3.628***	3.632***
	(0.097)	(0.093)	(0.098)	(0.096)
Distance ij		-0.107***		-0.064
		(0.032)		(0.042)
Same State ij			0.535***	0.385***
			(0.097)	(0.142)

Time Fixed Effects	Yes	Yes	Yes	Yes
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.*

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