The Roles of Geographic Distance and Technological Complexity in U.S. Interregional Co-patenting Over Almost Two Centuries

Milad Abbasiharofteh, Tom Broekel & Lars Mewes

Papers in Evolutionary Economic Geography

24.14

Utrecht University Human Geography and Planning

The Roles of Geographic Distance and Technological Complexity in U.S. Interregional Co-patenting Over Almost Two Centuries

Milad Abbasiharofteh
Department of Economic Geography, Faculty of Spatial Sciences, University of Groningen, the Netherlands.
m.abbasiharofteh@rug.nl (Corresponding Author)

Tom Broekel
University of Stavanger School of Business and Law, Norway.
tom.brokel@uis.no

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

Abstract

This paper examines how geographical proximity affected interregional co-patenting links in various technologies in the USA from 1836 to 2010. We classify technologies by their complexity and test whether that moderates the impact of distance on collaboration. Contrary to the 'death of distance' hypothesis, distance still matters for knowledge creation and exchange. Moreover, we show that the role of complexity has changed over time. However, this pattern reversed by the late 20th century, with collaborations in complex technologies becoming more resilient to distance than those in simpler technologies. However, this pattern reversed by the late 20th century, with collaborations in complex technologies becoming more resilient to distance than those in simpler technologies.

Keywords: network evolution, interregional collaboration, geographical proximity, technological complexity

JEL: O33, R12, N70, L14

1. Introduction

In the 1990s, prominent authors proclaimed the “end of geography” (O’Brien, 1992) and the “borderless world” (Ohmae, 1999) of modern economic activities. This eventually culminated in the “death of distance” (Cairncross, 2001) hypothesis, suggesting that geographic distance had become close to meaningless due to technological inventions substantially dropping communication and transport costs. The rise of telecommunication and the internet further fueled the idea that they had freed knowledge exchange of its historical spatial constraints (Friedman, 2005).

The rejection of this hypothesis has been widespread. The prominence of the proximity framework and relational approaches underscore the role of space in knowledge creation and exchange (Bathelt and Glückler, 2003; Boschma, 2005). Empirically, economic and innovation activities are shown to concentrate strongly in space, and this concentration seems to increase (Audretsch and Feldman, 1996; Florida et al., 2008; Sonn and Storper, 2008). The global distribution of knowledge is clearly “spiky” (Broekel et al., 2023; Florida, 2005), and the adverse effects of geographic distance on knowledge-related interaction and information diffusion are well documented (Abbasiharofteh and Broekel, 2020; Abbasiharofteh and Maghssudipour, 2024; Balland et al., 2013; Broekel, 2015; Makkonen et al., 2018).

However, knowledge domains are not homogeneous, and just looking at the average across all knowledge domains, as many studies do, may blur the differentiated effect of geographic distance. Researchers have started to investigate this issue. Most prominently, Balland and Rigby (2017) show that over the past 40 years, more complex knowledge domains have experienced stronger spatial diffusion constraints than simpler ones. Van der Wouden (2020) observed a similar pattern from 1836 to 1975. Yet, both studies differ in their conceptualization of knowledge-based interaction (citations vs. co-patenting), empirical approximation of complexity (method of reflection vs. NK model), and modeling of geographic distance (continuous vs. within/between-city interactions).

Consequently, we still do not know much about how geographic distance affects knowledge exchange in different technologies and levels of complexity. We also do not know if this effect has changed over time, as almost all knowledge domains have become more complex. This gap motivates the present paper, which examines how distance and complexity affect co-patenting activities in the USA from 1836 to 2010. We compared the association between distance and interregional collaborative ties across more than 600 technologies with different levels of complexity. We utilized gravity regression models and identified a stable and negative association between distance and collaboration despite the substantial advances in communication and transportation technologies.

Moreover, we observed significant variation across technologies related to their complexity. Interestingly, we found a reversal of the relationship between complexity and distance over time. In the 19th century, distance hindered the interregional diffusion of complex technologies more than it did in the case of “simple” ones. In the late 20th century, distance affected interregional knowledge exchange on complex technologies less. Many technologies even show a positive distance coefficient, perhaps suggesting that modern transport and communication technologies, economic incentives, and transformed organizational structure help overcome spatial barriers.

The paper's structure is as follows. Section 2 presents a brief review of the literature on knowledge network formation, particularly focusing on the proximity approach. Section 3 describes the patent database and the research method we employed. Section 4 presents and discusses the findings. Section 5 concludes the paper.

2. Knowledge exchange over space and time

In the 1990s, the rise of electronic communication technologies and the fall in communication costs led prominent authors to proclaim the "death of distance" (for a review, see Rietveld and Vickerman, 2003). This conjecture challenged the widely accepted view of geographical proximity as one of the most critical factors shaping the production and diffusion of knowledge in space and time. The relevance of geographical proximity lay at the heart of prominent concepts in national and regional innovation systems (Cooke et al., 1997), localized knowledge spillovers (Jaffe, 1989; Sonn and Storper, 2008), and (regional) knowledge networks (Bianchi and Bellini, 1991). Empirical works provide overwhelming support for the importance of geographical proximity alongside other types of proximity (Broekel and Boschma, 2012; Broekel and Hartog, 2013; Jaffe et al., 1993; Simensen and Abbasiharofteh, 2022; Ter Wal, 2014).

The scholarly debate did not end, as researchers developed additional hypotheses, qualifying the importance of geographic distance for innovation and knowledge-sharing activities. The first was that geographical proximity remained an obstacle to interaction and knowledge exchange, but its impact had become less relevant. Ground-breaking inventions improved transport and communication in the 19th and 20th centuries (Gordon, 2016; Taylor, 1989). Overland and overseas transport became significantly faster, more secure, and more reliable. In addition, the expansion of railway networks made more places accessible, such as the U.S. Midwest in the mid-19th century and the country west of the Mississippi River by the end of the 19th century (Atack et al., 2010).

In more recent decades, the possibilities of digital communication and air transport have further eased interactions and exchanges across space. Ter Wal (2014) tracked the development of the patent co-inventor network in the German biotechnology sector for more than 30 years. His results support this hypothesis; he found that the effect of geographical distance on knowledge exchange interactions decreased. Van der Wouden's (2020) study on U.S. co-inventor networks between 1836 and 1975 points in a similar direction.

In contrast, Balland et al. (2013), studying the video game industry covering 20 years, observed the formation of knowledge relations becoming increasingly sensitive to geographical distance. Menzel et al. (2017) explored the changing properties of the patent co-inventor network in the U.S. Research Triangle Park over 20 years. Their empirical results suggested a relatively stable influence of geographical distance over time. Consequently, how the effect of geographical distance on interaction and knowledge exchange has developed over time is unclear, motivating our empirical investigation. Similar to van der Wouden's (2019) findings in a longer and broader study, we expect that distance matters less for knowledge-based interaction in recent decades.

H1: The negative association between geographic distance and knowledge-based interaction diminishes over time.

However, the impact of geographic distance (and its change over time) may not be homogeneous across technologies. While technologies differ across multiple dimensions, recently, their complexity has received increasing attention (Balland et al., 2020; Balland and Rigby, 2016; Broekel, 2019; Mewes and Broekel, 2020). Due to the cumulative nature of knowledge, technological advancement and innovation have become increasingly harder to accomplish (Bloom et al., 2017). The creation of each generation of technologies and knowledge is based on what its predecessors have established (Aunger, 2010; Hidalgo, 2015; Howitt, 1999; Nelson & Winter, 1982). They also expand their range of functions. For example, "[d]*igital control systems [of aircraft engines] interact with and govern a larger (and increasing) number of engine components than [previous] hydromechanical ones"* (Prencipe, 2000, page 904). The tight integration of modern technologies into multi-technology systems adds further complexity, implying that the corresponding research and development (R&D) must cover and consider multiple technologies (Fai and Tunzelmann, 2001).

Advancing complex technologies requires more intensive collaboration than simple ones need (Balland and Rigby, 2016; Broekel, 2019; Hidalgo and Hausmann, 2009) because their inherently greater knowledge diversity necessitates interactions between specialized experts (Carbonell and Rodriguez, 2006; Pavitt, 1998). While collaboration has generally increased (e.g., Wuchty et al., 2007), the complex cases show the most pronounced growth in collaboration (Broekel, 2019; van der Wouden, 2020). Large knowledge shares of these technologies are generally tacit (Balland & Rigby, 2017). R&D processes related to complex technologies also involve greater uncertainty, ambiguity, failure, and experimentation (Broekel, 2019). Thus, having necessary competencies co-located or at least easily available is advantageous. Consequently, geographic distance can predictably constitute a greater obstacle for (joint) R&D in complex technologies than in simpler cases, which existing research also supports (Balland and Rigby, 2016; van der Wouden, 2018).

Moreover, the importance of such intra-region collaboration for complex technologies has intensified over time (van der Wouden, 2020). This finding supports the idea that geographic distance is a significant obstacle to knowledge interactions in more complex technologies. However, at the same time, the author showed that between 1835 and 1975, the percentage of inter-city collaboration rose from about 3% to almost 12.5%, and the average distance between co-inventors located in different cities grew from less than 200 km to more than 750 km (van der Wouden, 2020). Furthermore, he observed a non-linear association between complexity, geographic distance, and collaboration. For lower levels of complexity, the distance between inventors located in different cities increases before it seems to decrease for higher levels of complexity. Yet, he did not perform a rigorous empirical test of the related hypothesis. He also did not cover the time of advanced telecommunication and the internet, which motivates our paper's next two hypotheses. We follow his reasoning that developing more complex technologies promises more economic benefits (Mewes & Broekel, 2022). Consequently, complex technologies may offer more significant incentives to collaborate across greater distances. Mewes and Broekel (2022) point out that cross-city collaboration's benefits may only partially compensate for obstacles induced by high levels of complexity, translating into relatively lower inter-city collaboration. This implies that high levels of complexity should go hand in hand with geographical distance in constituting a significant obstacle to collaboration (Balland and Rigby, 2016; for similar arguments, see Sorenson et al., 2006).

Hence, the relationship between geographic distance and its influence on knowledge-based interactions will likely differ between simple and complex technologies that may have changed over time. More precisely, we argue that lowering communication and transportation costs may have impacted complex technologies more, for three reasons. First, developing complex technologies implies bringing together a greater diversity of expertise, usually unlikely to be found in just one location, and making interactions across space necessary. Second, expertise in complex technologies tends to be sparser and, consequently, more valuable (Kogut and Zander, 1992; Rivkin, 2000; Broekel, 2019). This specificity puts experts in these fields into a stronger negotiating position and reduces the incentives to move to places with an agglomeration of competence. Third, only a few places provide the conditions for competencies in complex activities to emerge and thrive (Hidalgo and Hausmann, 2009). If one requires such competencies, and one's location is outside such a place, collaboration (no matter the distance) may be the only possibility to access the capability one needs.

Consequently, due to technological developments, interacting across space becoming easier and cheaper has a greater impact on complex technologies. This is because they require this activity to a greater extent than simpler ones. Here is an extreme illustration: In the past, the high frequency of required interaction, combined with substantial costs associated with an individual interaction, implied very high absolute costs[1] that made distant collaboration for technology development in complex domains highly unlikely. The recent advancements in transportation and communication technologies have not reduced the need for frequent (face-to-face) interactions, but the costs associated with it have dropped substantially. Consequently, the absolute costs may have fallen to levels where such arrangements are attractive. In the case of simple technologies, such arrangements were always feasible because the required frequency of interaction was lower, implying that absolute costs might never have been prohibitively high.

H2: Geographic distance poses a greater challenge to knowledge-based interactions in complex technologies than in simpler ones.

H3: The negative association between geographic distance and the intensity of knowledge-based interactions utilizing complex technologies strengthens over time.

3. Empirical approach

3.1 Data

We relied on patent data to model knowledge generation and exchange. Analyzing patents (in addition to R&D alliance networks and scientific publications) has become one of the most straightforward methods for studying collaborative innovation processes (Cantner and Graf, 2006; e.g., Jaffe et al., 1993; Ter Wal, 2014). Co-patenting implies collaborations in the development of novel products or processes. These collaborative efforts usually entail substantial knowledge exchange and mutual learning (Ter Wal and Boschma, 2009) and, hence, meaningful knowledge (exchange) relations.

---

[1] Here, we use the term "absolute costs" in the sense of the product of frequency of interaction and costs of an individual interaction. In organizational decision-making, they matter as relative costs.

We used two patent databases: (1) HistPat (Petralia et al., 2016a), which covers historical patents that the United States Patent and Trademark Office (USPTO) registered from 1790 to 1975[2]; and (2) the USPTO digital database, which covers patents from 1975 onward (Li et al., 2014). These two databases are complementary, and users can merge them (Petralia et al., 2016b). The result provides a database covering more than 170 years of data, with information on the patent's grant date, its technological classes, and its places of origin. We used 3,220 Federal Information Processing Standards (4-digit FIPS) as unique identifiers of U.S. counties, to systematize the places of origin[3]. However, whether these locations refer to the applicant or the inventor is unknown. Unlike modern patent data, geographic locations are not inventor- or applicant-specific but a mix of both. This issue was unlikely to be a problem in earlier years when applicants and inventors were likely identical because most patents were granted to individuals. In later years, this changed with an increasing share of patents granted to firms, implying that inventor and application locations do not necessarily coincide geographically. Lacking a way to identify inventors and applications, we could not resolve this drawback at this time, and it remains a major limitation of the present study.

Moreover, as we could not differentiate distinct actors located within regions, we had to focus exclusively on interregional relations. Therefore, we had to follow an established practice in the field (Broekel and Hartog, 2013; Hoekman et al., 2009), modeling proximities at the regional level, despite the underlying rationale at the individual and organizational levels.

We chose FIPS as the spatial units of our analysis, implying that they constitute the network nodes, with links being different counties co-occurring on patents. For each decade, we aggregated all patents with an application date falling within the period (*t-9* to *t*), generating 18 non-overlapping time-windows (between 1830 and 2010) of U.S. interregional co-inventions. For the sake of simplicity, we refer to each time-window with the most recent year in the corresponding time-window. For instance, time-window 2010 refers to inventive activities between 2001 and 2010.

Figure 1 provides an overview of the spatial distribution of interregional collaborative ties and their intensity over time. In the 19th century, interregional inventive activities primarily concentrated on the East Coast. Scholars acknowledge the early 20th century as a time of significant transformation in the structure and efficiency of the United States' innovation economy (Lamoreaux and Sokoloff, 1996). In the early 20th century, the Midwest region contributed significantly to the co-inventive network, and collaborative ties between the East and West Coasts started emerging. Teaford (1994) explains the decline of the Midwest from the 1950s onward, where previously prosperous cities, such as Detroit, Chicago, and Cleveland, became known as the "Rust Belt."

In contrast, California experienced population growth and emerged as the most populous U.S. state since the 1960s. Toward the end of the 20th century, intensified interregional relations between California, Texas, and Washington regions and regions on the East Coast became visible. In contrast, connections between Midwestern regions and others decreased.

Figure 2 depicts the number of patents, joint patents, and interregional collaborative ties. All three variables show an overall increase, with a stagnation period in the early- and mid-20th

[2] The examination of patent applications was implemented in the 1830s. Thus, we studied patents from 1830 onward.

[3] For consistency purposes, we refer to county FIPS as regions. Accordingly, collaborative ties that cross FIPS boundaries denote interregional relations.

century, due to the impact of World Wars and the Great Depression. Notably, Figure 2 shows that inventive activities increasingly required joint work (Wuchty et al., 2007), most likely due to their growing complexity (Broekel, 2019; van der Wouden, 2020).

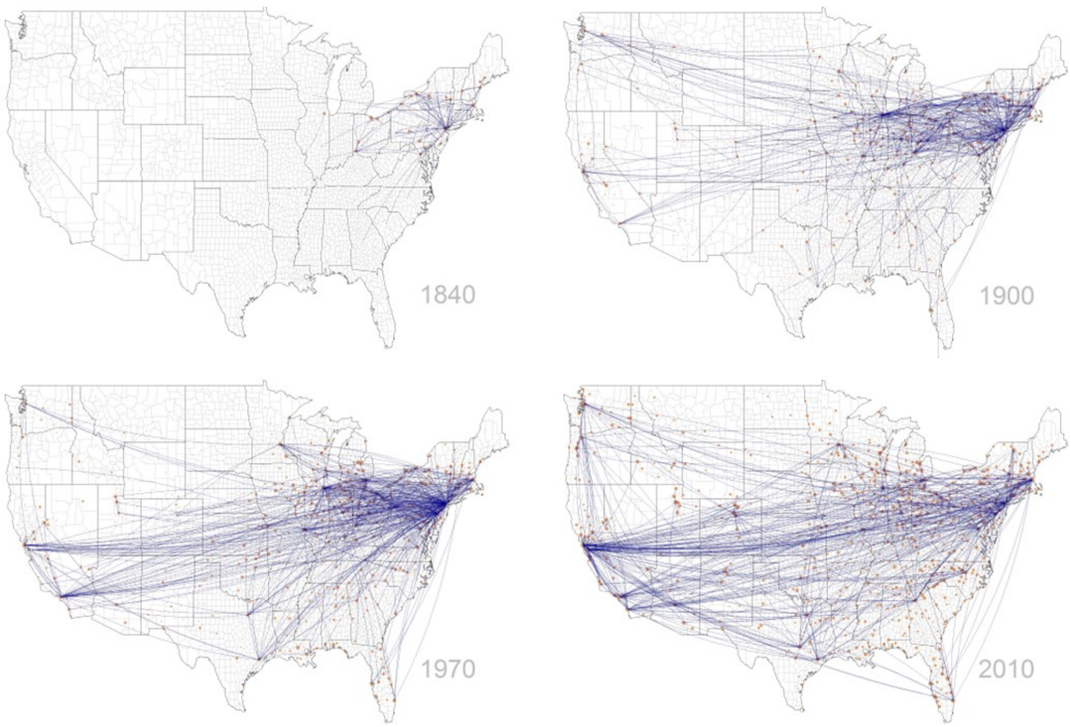

Figure 1. Interregional co-invention network in four time-windows.
The dots correspond to the number of patents in each county, and the lines represent the 600 most-repeated interregional collaborations (the first map demonstrates all collaborations in the 1840 time-window).
Source: Authors' illustration based on HistPat (Petralia et al., 2016a) and USPTO digital database (Li et al., 2014).

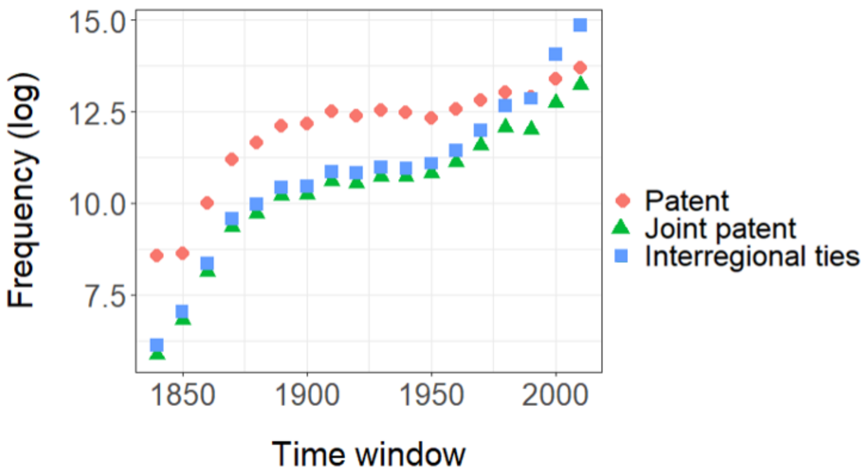

Figure 2. The number of patents, joint patents, and interregional collaborative ties (region-pairs) between 1840 and 2010.
Note: This graph is based on 10-year time-windows. A joint patent is a patent granted to two or more inventors who have collaborated on the invention.
Source: Authors' illustration based on HistPat (Petralia et al., 2016a) and USPTO digital database (Li et al., 2014).

Alongside the shifting geography, the USA experienced continuous growth of its transportation system, which frequently did not "just" replace existing connections but actually "*put new locations* [like San Francisco] *on the map*" (Atack, 2017). Figure 3 illustrates this continuous expansion in the case of the railroad system. Train connections fostered interregional trade and people's mobility, thereby facilitating spatial knowledge diffusion and exchange. Not surprisingly, the first significant increase in East Coast–West Coast patent collaborations in U.S. history (from about 50 in 1860 to about 300 in 1880) occurred when the first railways connected the two coasts. Similar patterns emerged with the establishment of regular air transport connections and low-cost carriers. Air transport altered the ranking of places' reachability and likely reduced and shuffled the effect of geographical proximity. For instance, at the end of the 1930s, New Yorkers could reach Los Angeles by plane in 18 hours but had to travel over land for a full day to reach Philadelphia (Paullin, 1932).

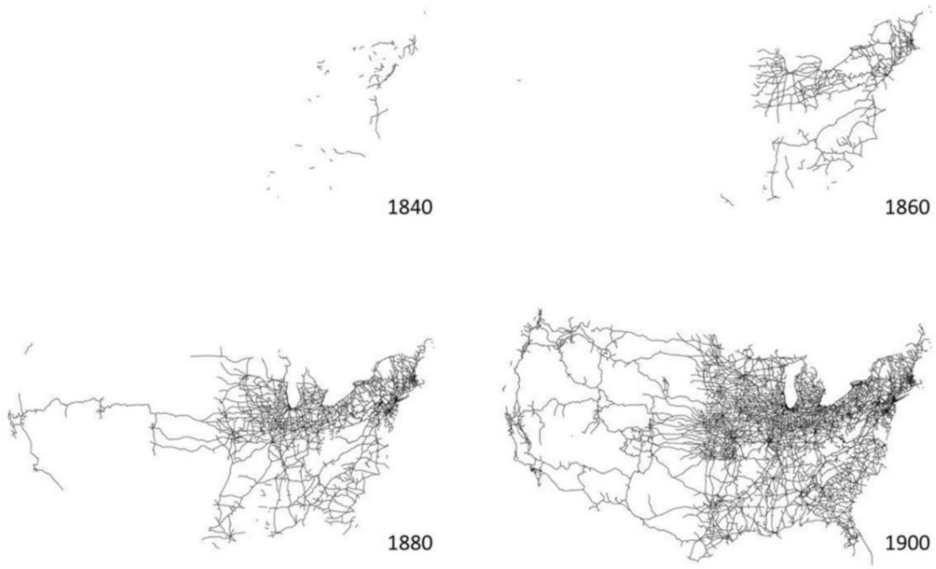

Figure 3. The expansion of the railroad system in the USA.
Source: Authors' illustration based on Atack (2017).

3.2 Variables

Dependent variables

The unit of analysis is region-pair potential relations. Given the number of regions (*n=3,220*) and technologies (*c=654*), this results in 3,389,413,860 potential region-pairs-technologies *[(n×n-1)/2]×c*, where *n* denotes the number of regions and *c* the number of technologies. Inventive interregional activities occurred between relatively few regions in the past 170 years (the mean of realized interregional relations: 15,421 of more than 3 billion potential ones), leading to zero-overdispersion, i.e., a dominating share of zeros. This issue well-known in trade studies. Prominently, Burger et al. (2009) argue that "*a distinction can be made between pairs of countries with exactly zero probability of trade, pairs of countries with a non-zero trade probability that still happen not to be trading in a given year, and pairs of countries that are trading*" (page 175). Translating this to our situation, a crucial differentiation of zeros, representing the absence of interregional patenting between two regions that have zero probability for such type of patenting, and zeros indicating the absence of a link (although, theoretically, a positive probability of such a link could exist). We argue that when regions are not part of any interregional patenting at a given time, this indicates the absence of the general capability of such activities. For instance, in the early periods, many regions lacked any (patenting-related) inventive activities, a prerequisite for interregional patenting. But even in later periods, there are regions with documented inventive activities but no interregional linkages. We interpret both cases as indications that basic conditions for interregional collaboration were unmet (e.g., insufficient infrastructure). To avoid this (directly) unobservable factor biasing our results, we restricted the sample of observations to those regions with at least one interregional tie in a given technology within each 10-year time-window. The minimum of one existing interregional tie thereby represents the documented ability of regional inventors to engage in interregional patenting. As a result, the number of observations substantially decreases and dynamically changes over time.

On this basis, we construct our dependent variable as a dummy variable that takes the value of one if at least one collaborative tie is recorded for a region-pair in a given technology, and zero otherwise.

Figure 4 shows that the number of observations (region-pairs across all technologies) substantially varies over time, generally tending to increase until WWI. Subsequently, they decreased before they continued growing after WWII. The figure clearly shows the disruptive effect of the World Wars and the period in between.

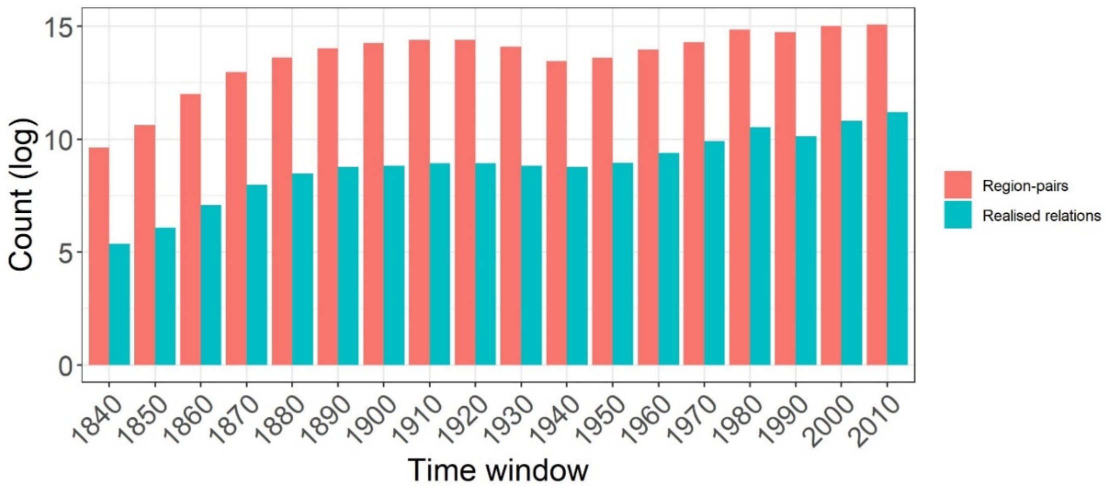

Figure 4. The number of region-pairs (including regions with at least one interregional relation) and realized interregional relations.
Note: Realized interregional relations correspond to the number of region-pairs whose inventors collaborated to develop at least one invention. The maximum number of realized interregional relations is equal to that of region-pairs.

We observed a positive correlation between the population of regions and the diversity of their technological portfolios, which the entropy index approximated (the Pearson correlation coefficient: 0.74). However, the correlation between population and the number of interregional ties is relatively weak (0.19). Furthermore, there is no substantial correlation between the diversity of technological portfolios and the number of interregional relations (0.1). Conversely, there is a positive correlation between the number of patent applicants in each region and the number of collaborative interregional ties (0.69).

Relational explanatory variables

The core explanatory variable is geographic distance. We defined *DIST* by taking the logarithm of the Euclidean distance between regions in kilometers. As an alternative, we created the variable *STATE*, taking a value of one when two regions are located within the same U.S. state, and zero otherwise. The second variable, however, has two interpretations, as it captures both geographic and institutional distance (differences in state-specific regulations and legal frameworks). Consequently, it primarily functions as a robustness test. In contrast to the other variables that appear below, the values of these two variables are time-invariant.

Relational control variable

We followed the literature in approximating the "cognitive" distance between two regions using their "technological" distances[4] as the Spearman rank correlation of the patent portfolios of regions represented them (Abbasiharofteh, Kogler et al., 2023). The correlation coefficients indicate how similar interacting regions are, regarding their technological schemes. We multiplied the variable by minus one to express the result in terms of distance (*TECH*)[5]. Although the number of technologies and regions involved in co-inventorship increases, Figure 5 shows that the average technological distance between collaborative regions does not substantially vary over time. More interestingly, Figure 5 suggests while the interquartile ranges of distance between collaborative regions overlap across most time-windows, the median and maximum values of geographic distance show a four-fold increase. For instance, in 1840, the maximum distance of interregional relations was less than 2,000 kilometers. In the last time-window (2010), we observed a collaboration between inventors from Honolulu County (Hawaii) and Newport County (Rhode Island), more than 8,000 kilometers apart. This observation aligns with this section's description of the evolution of economic and co-inventive activities in the 19th and 20th centuries.

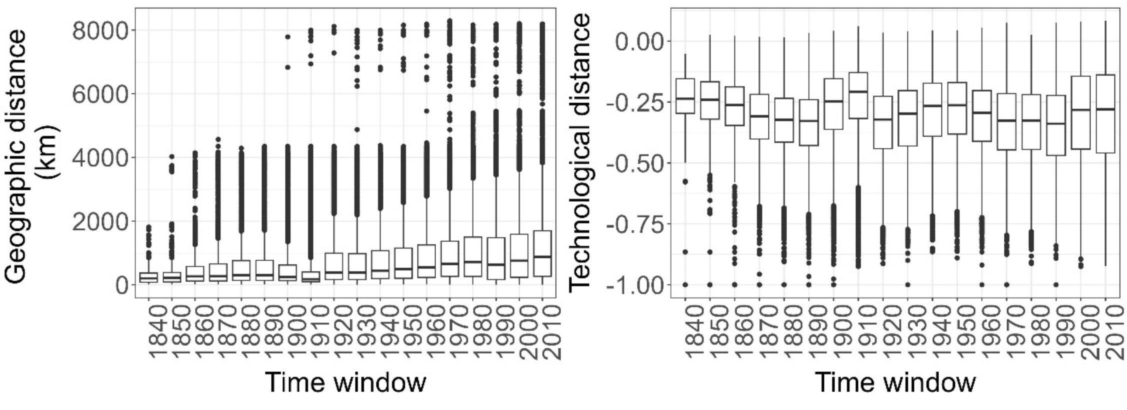

Figure 5. The geographic and technological distance of interregional co-inventive relations.

Technological complexity

To approximate technological complexity, we relied on the measure of structural diversity that Broekel (2019) proposed. The measure captures the heterogeneity of relational structures

[4] Patent portfolios are created based on the CPC codes at the 4-digit level.
[5] To avoid confusion, we express all variables in terms of distance (not proximity).

among technologies' (knowledge) elements, the so-called combinational network representation of technologies. The greater this heterogeneity is, the larger the information content of technologies' combinatorial networks is, and the more complex the associated technology is. Broekel (2019) applied this approach to estimate the complexity of patent classes (as an approximation of technologies). For each 4-digit technology class (CPC[6]), he constructed an individual combinatorial network based on 10-digit CPC codes' co-occurrence on patent documents. Their structural diversity is measured with the Network Density Score (NDS), created by Emmert-Streib and Dehmer (2012).

To assign complexity scores to technologies, we calculated the technological complexity of patents associated with a specific 4-digit CPC code, following the procedure Broekel described (2019).[7] Then, we averaged the annual values for each technology in each time-window, to avoid potential biases that annual patent number fluctuations introduced[8].

The complexity indices suggest that technologies associated with *Machines or engines for liquids* (CPC code: F03B) were the most complex, and *Book covers* (B42D) were the least complex technologies in 1840. In 2010, *Ship-lifting devices or mechanisms* (E02C) and *the Propulsion of electrically-propelled vehicles* (B60L) were the least and the most complex technologies, respectively. Figure 6 provides descriptive statistics for the obtained complexity scores of technologies. Importantly, when a new technological code is added to the previous ones, the USPTO redefines all existing patent classes (e.g., CPC codes) (Petralia, 2020). This fact implies that the number of technologies (with non-zero complexity scores) varies across time-windows. The variable *COMPLEXITY* represents a time-varying continuous value at the level of technologies. The difference between the units of analysis (i.e., region-pairs and technologies) calls for a specific empirical approach that the next section presents.

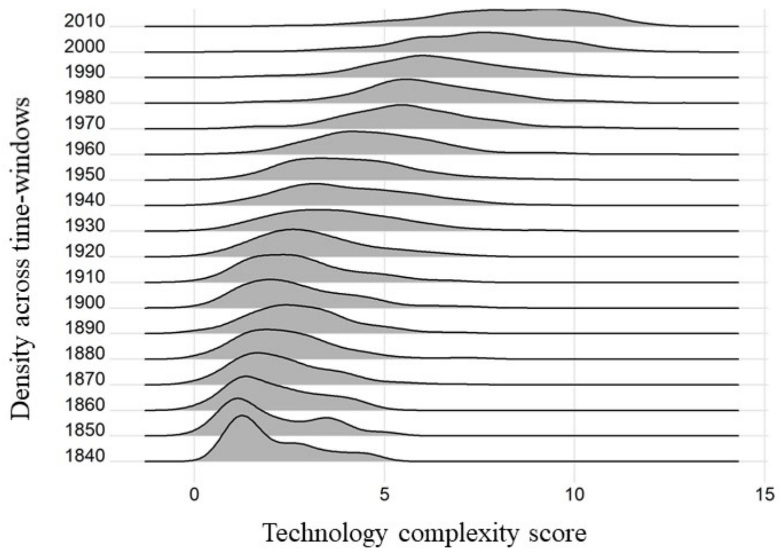

Figure 6. Kernel density estimate of technology complexity scores across 18 time-windows.

---

[6] The Cooperative Patent Classification (CPC) describes technologies and comprises nine primary sections, namely A-H and Y, which further branch out into classes, sub-classes, groups, and sub-groups. Within this classification system, around 250,000 entries exist, encompassing various classifications.

[7] We opted for the CPC classification instead of the U.S. patent classification, to remain comparable to existing studies using this measure in the European context.

[8] To minimize the effect of outliers, we also used median instead of mean. The results suggest no significant difference between mean and median values.

Note: Technology complexity scores quantify the diversity of relational patterns within the constituent sub-classes of technologies, represented as combinational networks. Higher levels of heterogeneity indicate greater informational content within the networks and correspondingly increased complexity of the associated technology. Broekel (2019) utilized this methodology to assess the complexity of patent technological classes (CPC codes).

3.3 Estimation strategy

We employed a spatial-interaction model (also known as the gravity model) to test the hypotheses. In the 1960s, this model became popular for analyzing trade between countries (e.g., Hasson and Tinbergen, 1966). We followed recent applications in using the model to account for the establishment and intensity of interregional co-invention (Maggioni et al., 2007; Scherngell and Barber, 2011)[9]. The model is designed to deal with relational observations (i.e., region-pairs). It features variables at the region-pair levels (e.g., geographic distance between two regions).

Although researchers often use logit models to estimate models with binary dependent variables, we utilized Linear Probability Models (LPM) for two reasons. First, including dummy variables as fixed effects in a non-linear model can lead to biased results (Gomila, 2021). This problem motivates using LPM, which the relatedness literature has widely used to model the entry of regions into new economic activities (for instance, see Boschma et al., 2023). Second, given the size of the data and the number of fixed effects, conducting logistic regressions is prohibitively expensive due to computational time and resource requirements[10].

Notably, the model includes both time-variant and time-invariant explanatory variables. In other words, the variables *DIST* and *STATE* do not change across 18 time-windows, whereas *TECH* varies over time for a given region-pair.

To test the first hypothesis, we operationalized the spatial-interaction model by estimating LPM, where $P_{ij,t,c}$ denotes the probability of at least one collaborative tie between inventors located in regions i and j in time-window t utilizing technology c. The LPM model is defined as follows:

$$P_{ij,t,c} = \beta + \Gamma DIST_{ij} + \Lambda STATE_{ij} + \Theta TECH_{ij,t} + \omega_c + \xi_i + \alpha_j + \varepsilon_{ij,t,c} \qquad (1)$$

To control for unobserved heterogeneities, we include ω_c, ξ_i, and α_j, which are technology (CPC codes), region i, and region j fixed effects, respectively. The error term is $\varepsilon_{ij,t,c}$.

Equation 1 does not feature the complexity dimension because the (time-variant) variable *COMPLEXITY* is undefined at the regional and the relational level. Therefore, we opted for a two-stage approach to test the second and third hypotheses. In the first step, we estimated an individual interaction model for each time-window.

[9] The main reason for not using dynamic network analysis tools (e.g., stochastic actor-oriented models) is difficulties in fitting and converging models entailing millions of relations.

[10] For estimating LPM models, we used a high-performance computing cluster and allocated three terabytes of RAM.

$$P_{ij,t,c} = \Upsilon + \Phi DIST_{ij} \times CPC + \Psi STATE_{ij} + \Omega TECH_{ij,t} + \omega_c + \xi_i + \alpha_j + \varepsilon_{ij,t,c} \qquad (2)$$

The coefficients of $DIST_{ij} \times CPC$ are the basis for the second step, the meta-regressions. We investigated the interplay between the effect size of $DIST_{ij} \times CPC$ coefficients and the complexity of technologies with a meta-regression analysis. This method is "the regression analysis of regression analyses" (Stanley and Jarrell, 1989, page 161). Medical and psychological research have frequently used it, but increasingly it also finds applications in regional and urban studies (Groot et al., 2007; Melo et al., 2009).

We estimated mixed-effects meta-regression models using a residual maximum likelihood approach (REML) (Viechtbauer et al., 2015). Various studies have widely used the mixed-effects approach, due to its flexibility in allowing for within- and between-period variations (Benjamini, 2020).

We estimated meta-regressions with *COMPLEXITY* as the main predictor and the effect sizes of interest as the dependent variable (coefficients of $DIST_{ij} \times CPC$ denoted as $y_{k,c}$ below). The regression coefficients indicate whether the associations with geographic distance are conditioned on *COMPLEXITY* and, if so, the extent to which these associations change over time. We denote the meta-regression model as follows:

$$y_{k,c} = O + \Xi COMPLEXITY_c + XPatentNumber_c + W_{k,c}Y + \varepsilon_{k,c} \qquad (3)$$

where $y_{k,c}$ is the meta-dependent variable corresponding to the effect size of $DIST_{ij} \times CPC$ coefficients that the k^{th} *model*, which estimates the association of interregional tie formation in technology c (i.e., CPC codes), reports. The meta-independent variable $COMPLEXITY_c$ captures the technological complexity of technology c in a given time-window. We control each technology's size (the logarithm of the number of patents) in each time-window (i.e., $PatentNumber_c$). The random effects capture between-studies variability, denoted as $W_{k,c}Y$. In our case, it is equivalent to the between-time-period variability since the individual "studies" are the individual estimations for each time-window. The coefficient of this regression model expresses the correlation between the coefficients of geographic distance and the degree of complexity of the underlying technology[11].

4. Results and discussion

The preceding section describes the number of observations in each time-window reflecting the extent to which inventors in various regions engage in interregional co-inventive activities. Notably, a negative correlation exists between the number of observations and the goodness of fit of the estimated models (Figure 7). A probable reason for this observation is that the number of potential relations grows much more than that of the realized relations, as the co-inventorship network features more regions (see Figure 4). This inflation of the number of zeros in the dependent variable (no ties between pairs of regions) reduces the goodness of fit[12]. We visualize

[11] To estimate meta-regressions, we used the R package *metafor* by Viechtbauer (2010).
[12] The observation number and the goodness of fit of estimated models with interaction terms resemble the ones the manuscript presents.

the coefficients of the variables and their upper and lower margins of error, to trace their development over time.

The first interesting finding concerns one of our control variables, namely, technological distance. As the continuously negative coefficient shows (left panel in Figure 8), initiating collaboration between two technologically distant regions is less likely than for two proximate regions. This finding aligns with similar findings for European regions (Morescalchi et al., 2015) and contrasts with that of other studies, such as Lata et al. (2017). Given the different empirical approaches and data sources these studies used, finding the source of this inconclusiveness would require further investigations. Nevertheless, our investigation adds evidence that technological distance has negatively influenced the likelihood of spatial interaction for a long and continuous time.

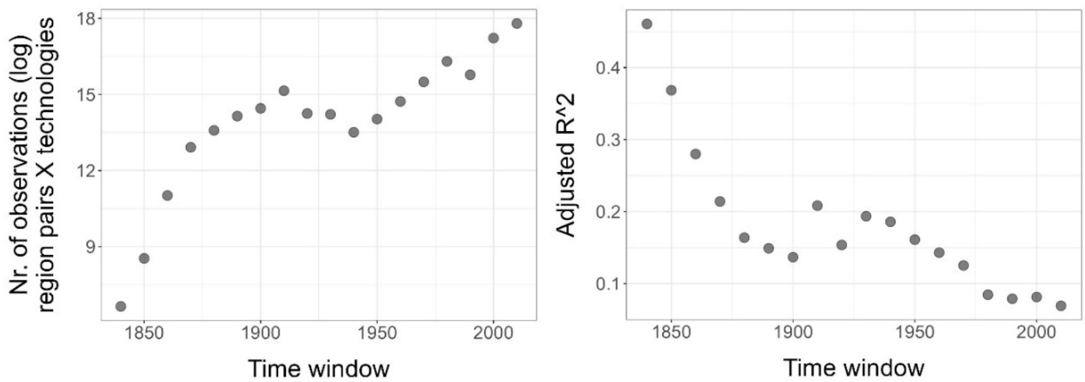

Figure 7. The number of observations (left) and models' goodness of fit across time-windows and 654 technologies (right).

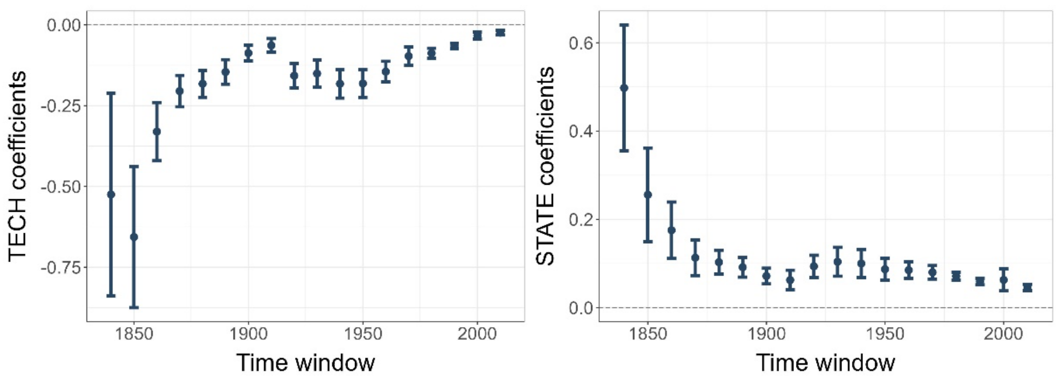

Figure 8. The estimated coefficients for technological distance (*TECH*) and co-location of regions in the same state (*STATE*) with the corresponding margin of error.

We now turn toward our main research interest, namely, the role of geographical distance. Figure 8 [right] features the results concerning inventors' co-location in the same U.S. state (*STATE*). We observe significantly positive coefficients for this variable in all periods. Co-location in the same state increases the likelihood of collaborative tie formation. Yet, the magnitude of the coefficients is somewhat lower in recent years than those in the late 19th century, suggesting that in the past, being in the same state was of greater importance for establishing interregional collaborations. However, this variable has a less clear interpretation than geographic distance because it might also capture institutional proximity. Therefore, the remainder of the paper focuses on the results for geographic distance (Figure 9), primarily using *STATE* as a control for institutional distance.

Figure 9 illustrates the coefficients obtained for the geographic distance variable (*DIST*). All coefficients except those of the first two time-windows are significantly negative, implying that we can clearly reject Hypothesis **H1**. Geographical distance has not become less of an obstacle for interregional inventor collaborative relations. Two robustness checks substantiate this finding (see Appendix A).

Potentially, this may be due to transportation costs not having fallen much further and remaining relatively stable in more recent years, at least concerning airfares, the most relevant mode of transportation for inventors in the USA in recent times (Perry, 2014).

Van der Wouden's (2020) study also examines the historical aspect of inventors' collaboration on U.S. patents from 1836 to 1975. His finding demonstrates that patenting has increasingly become a collaborative work, concentrating in large cities. Similarly, the study by Lin et al. (2022) suggests that patents with inventors situated in the same city are more likely to be disruptive. Specifically, these patents are more susceptible to receiving citations that do not refer to prior works. Our results complement the findings of van der Wouden (2020) and Lin et al. (2022) by showing that the collaboration-facilitating role of geographic proximity does not stop at city boundaries. Even beyond these, location near other inventors still increases chances of collaboration.

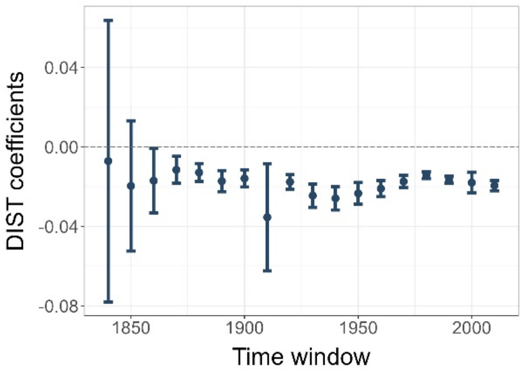

Figure 9. The estimated coefficients for geographic distance (*DIST*) with the corresponding margin of errors.

The result raises the question about the cause of this development. That is, why does geographic distance impede cross-region relations, despite all improvements in transportation and communication that have excelled in reducing interactions across space? As the theory section discusses, one reason might be the growing technological complexity that requires more intense interactions facilitated by geographical proximity (Broekel, 2019; Balland and Rigby, 2016). In line with our hypotheses, we expect that the coefficients of the geographic distance variable *DIST* relate negatively to technologies' complexity in each period (H2), and the relationship grows stronger over time, implying that the slope of this relationship becomes steeper over time (H3). We tested the two hypotheses, employing the second-stage meta-regression. The result shows the association between coefficients of the interaction term (i.e., the geographic distance variable multiplied by technology dummies: $DIST_{ij} \times CPC$) and the levels of technological complexity (Table 1).

Our results show that the relationship between the complexity and the geographic distance of interregional collaborations varies in both sign and magnitude. The hampering effect of geographic distance on distant collaboration between inventors was larger for simple than for complex technologies. Accordingly, H2 is only confirmed for the 19th century. For the 20th century, we must reject it since we find that the opposite holds true. The distances of interregional collaboration tend to be larger for complex technologies than for simpler ones. The latter also implies that we must reject H3. Over time, geographic distance has lost its hampering effect on collaboration for complex technologies, the opposite of our expectations.

The change in the sign of the complexity variable in the meta-regression between the 19th and the early and mid-20th centuries likely relates to the fundamental shift in inventive activities in the U.S. during the Second Industrial Revolution. That revolution shifted the nation's economy from agriculture to manufacturing and services. The analysis by Hartog et al. (2022) of inventive activities between 1850 and 1940 highlights engineers taking over a pivotal role in innovation, and inventions becoming collaborative teamwork. Moreover, new organizational structures, such as R&D labs, enabled connections between inventors across various locations.

Esposito (2023) and van der Wouden (2022) confirmed the growing significance of large collaborative teams at the local level, most prominent in disruptive technological domains. Our findings suggest that the high demands of complex technologies, in terms of complementary competencies and knowledge and the higher economic returns, extend these collaboration activities beyond regional boundaries. As competencies associated with complex technologies are spatially less ubiquitous (Hidalgo and Hausmann, 2009) it seems likely that, in many cases, they will not be in the same location, forcing inventors to collaborate across multiple (distant) locations. Add to this the economic benefits of complex technologies that, in contrast to the smaller ones of simple technologies, outweigh the costs of overcoming larger geographic distances. Given the change in the coefficient's sign in the second half of the 20th century, we suspect that specifically cheap interstate air travel and improved communication technologies in the 1970s contributed to this development. For instance, Southwest Airlines started interstate service in 1979, whereby inventors could meet face-to-face to troubleshoot and discuss complex technical issues.

Our findings resonate with those of Esposito (2023). He highlights that since 1975, distant inventors located in U.S. cities with diverse technological domains would more likely introduce high-impact patents, thereby facilitating an unconventional combination of technologies. While Van der Wouden (2020) reported the opposite. However, he considered intra-regional collaboration. Therefore, it seems plausible that the lower impact of geographic distance on collaboration in the case of complex technologies is restricted to the scenarios of long-distance collaboration[13].

	<1900	**<1940**	**<1980**	**>1980**	**All**
intercept	-0.168 (0.006)***	-0.117 (0.007)***	-0.025 (0.005)***	-0.046 (0.004)***	-0.086 (0.004)***
PatentNumber	0.023 (0.001)***	0.014 (0.000)***	0.004 (0.000)***	0.004 (0.000)***	0.010 (0.000)***
COMPLEXITY	-0.001 (0.000)***	0.000 (0.000)	0.000 (0.000)***	0.001 (0.000)***	-0.000 (0.000)

[13] To ensure the robustness of our models, we created several controls for regions. We considered the size of regions, approximated by population counts. We obtained the data on the regional population between 1836 and 2010 from the published records and the U.S. Census Bureau website. We also captured regions' inventive activities by the number of patents assigned to region-pairs. Moreover, we estimated the Gini coefficients of the technological portfolios of regions (4-digit CPC codes) in each 10-year time-window. However, our models having two sets of region-fixed effects caused all these region-level controls to drop out.

	<1900	<1940	<1980	>1980	All
tau.squared	0.000	0.000	0.000	0.000	0.000
cochran.qe	6988.705	6736.672	6283.218	39485.667	167085.744
p.value.cochran.qe	0.000	0.000	0.000	0.000	0.000
cochran.qm	1544.297	1854.526	955.424	315.735	3179.434
p.value.cochran.qm	0.000	0.000	0.000	0.000	0.000
DF Resid.	1625	1840	2223	1753	7450
Log Likelihood	3710.986	4267.336	6925.477	4415.608	18193.167
Deviance	-7421.971	-8534.672	-13850.954	-8831.215	-36386.334
AIC	-7411.971	-8524.672	-13840.954	-8821.215	-36376.334
BIC	-7385.005	-8497.085	-13812.421	-8793.870	-36341.755
AICc	-7411.934	-8524.640	-13840.927	-8821.181	-36376.326
nobs	1628	1843	2226	1756	7453

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 1. The meta-regression shows the relation between complexity and the coefficients of the geographic distance across technologies.

5. Conclusion

Many studies have analyzed the role of geographic distance in forming and developing knowledge-based interactions. Prominently, Friedman (2005) argued that major political changes (e.g., the fall of the Berlin Wall) and technological advances (e.g., the invention of the internet) had created forces that “flattened” the world, implying that geographic distance had lost (or, at least, was losing) its relevance as an obstacle to interaction and collaboration (see also the “death of distance” debate). This conjecture stimulated the emergence of a rich literature seeking to understand if and how the effect of geographic distance changes over time and across various contexts (Balland et al., 2013; Broekel, 2015; Ter Wal, 2014).

However, until recently, this discussion and the literature have rarely considered the heterogeneity among technologies, which may mediate the relevance of geographic distance for knowledge-based interactions. Balland and Rigby (2017) prominently argued that higher levels of technological complexity might make geographic distance a more significant obstacle to technological diffusion in space. Van der Wouden (2020) investigated this argument for 1836 and 1975, observing complex technologies relying strongly on intra-city collaboration.

Our study adds to this literature by investigating the evolution of the interregional co-inventor network of U.S. regions between 1836 and 2010. Its focus was on the (changing) relationship between geographic distance and the formation of interregional collaboration. Our findings clearly show that geographical distance, at least in interregional collaboration over the last 170 years, has kept its importance as obstacle, despite all advances in transport and communication technologies. The empirical results also challenge previous findings in the literature (Balland and Rigby, 2017; Van der Wouden, 2020) by suggesting that higher levels of technological complexity do not make distance a greater obstacle to collaboration, at least in the second half of the 20th century. We conjecture that this phenomenon arises due to the inherent knowledge requirements of complex technologies that require drawing upon a broader range of expertise and offering significant economic gains through collaboration. Additionally, the advancements in transportation and communication infrastructure have significantly reduced the costs of knowledge-based interactions, making such collaborations more feasible and lucrative for complex technologies than for simpler ones. Importantly, the precise reasons for this observation remain speculative and require more investigations in the future.

Having reviewed our study's main findings, we acknowledge our research's limitations. Most importantly, we lack information on intra-regional knowledge-based interactions. Our analysis relies on patent data, known to reflect only a fraction of innovation and collaboration activities. However, this drawback was unavoidable; no other available data source covers such time spans.

Another issue that future studies might overcome is the aggregation at the regional level. Currently, no information on inventors or applicants and the spatial structure of multi-locational firms is available for the HistPat database, which restricted us from exploring collaboration patterns between regions. This issue introduces the danger of ecological fallacy and neglecting crucial organizational heterogeneity. We encourage researchers to replicate our study using recently published firm-level data, such as trademark and web-text and hyperlink data (Abbasiharofteh et al., 2022; Abbasiharofteh, Kinne et al., 2023; Abbasiharofteh, Krüger et al., 2023; Nathan and Rosso, 2015). Similarly, we acknowledge that county-specific regulations and legal frameworks may serve as catalysts or obstacles to interregional relations. Due to the temporal scope of our investigation spanning 170 years, we could not ascertain the level of institutional proximity at the county level.

Nevertheless, our findings raise questions for future research. For instance, the USA experienced several waves of immigration over the last 200 years. Empirical studies suggest that regions in which most immigrants settled during the 19th and the early 20th centuries have enjoyed economic prosperity in the long run (Rodríguez-Pose and Berlepsch, 2015). In this light, studying the impact of immigration on knowledge diffusion and interaction will improve our understanding of how sociocultural diversity influences spatial interactions and knowledge diffusion (Diodato et al., 2022).

Moreover, Sokoloff's (1988) findings on the impact of railways on the innovative activities of distant areas in the early 19th century call for an examination of the influence of roads and air transport development on interregional collaborations, to better understand how and why interregional collaborations emerge and intensify.

Our empirical results are not only crucial for advancing our understanding of the interplay of technological complexity, collaboration, and geography. They are also essential for the sectoral and mission innovation policy design, which increasingly seeks to stimulate collaboration (Broekel and Graf, 2012; Janssen and Abbasiharofteh, 2022). Our empirical findings shed some light on whether to support local or interregional collaboration in different technological contexts, which motivates further research today and, we hope, informs new policies tomorrow.

References

Abbasiharofteh M, Broekel T, 2020, 'Still in the shadow of the wall? The case of the Berlin biotechnology cluster' Environment and Planning A: Economy and Space 53(1) 73-94.

Abbasiharofteh M, Castaldi C, Petralia S G, 2022 From patents to trademarks: towards a concordance map (European Patent Office (EPO)).

Abbasiharofteh M, Kinne J, Krüger M F C, 2023 Leveraging the digital layer. The strength of weak and strong ties in bridging geographic and cognitive distances.

Abbasiharofteh M, Kogler D F, Lengyel B, 2023, 'Atypical combinations of technologies in regional co-inventor networks' Research Policy 52(10) 104886.

Abbasiharofteh M, Krüger M, Kinne J, Lenz D, Resch B, 2023, 'The Digital Layer: Alternative Data for Regional and Innovation Studies' Spatial Economic Analysis 1–23.

Abbasiharofteh M, Maghssudipour A, 2024, 'Driving Forces Behind Relational Knowledge Sourcing in Clusters: Single- and Multilevel Approaches' Journal of the Knowledge Economy 28(3) 1.

Atack J, 2017, "Historical transportation of navigable rivers, canals, and railroads in the United States" (ICPSR - Interuniversity Consortium for Political and Social Research).

Atack J, Bateman F, Haines M, Margo R A, 2010, 'Did Railroads Induce or Follow Economic Growth?' Social Science History 34(02) 171–197.

Audretsch D, Feldman M P, 1996, 'R&D spillovers and the geography of innovation and production' American Economic Review (86) 630–640.

Balland P-A, Jara-Figueroa C, Petralia S G, Steijn M P A, Rigby D L, Hidalgo C A, 2020, 'Complex economic activities concentrate in large cities' Nature human behaviour.

Balland P-A, Rigby D, 2016, 'The Geography of Complex Knowledge' Economic Geography 93(1) 1–23.

Balland P-A, Vaan M de, Boschma R, 2013, 'The dynamics of interfirm networks along the industry life cycle: The case of the global video game industry, 1987-2007' Journal of Economic Geography 13(5) 741–765.

Bathelt H, Glückler J, 2003, 'Toward a relational economic geography' Journal of Economic Geography 3(2) 117–144.

Benjamini Y, 2020, 'Selective Inference: The Silent Killer of Replicability' Harvard Data Science Review 2(4).

Bianchi P, Bellini N, 1991, 'Public policies for local networks of innovators' Research Policy 20(5) 487–497.

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

Boschma R, Miguelez E, Moreno R, Ocampo-Corrales D B, 2023, 'The Role of Relatedness and Unrelatedness for the Geography of Technological Breakthroughs in Europe' Economic Geography 99(2) 117–139.

Broekel T, 2015, 'The Co-evolution of Proximities – A Network Level Study' Regional Studies 49(6) 921–935.

Broekel T, 2019, 'Using structural diversity to measure the complexity of technologies' PloS one 14(5) e0216856.

Broekel T, Boschma R, 2012, 'Knowledge networks in the Dutch aviation industry: The proximity paradox' Journal of Economic Geography 12(2) 409–433.

Broekel T, Graf H, 2012, 'Public research intensity and the structure of German R&D networks: A comparison of 10 technologies' Economics of Innovation and New Technology 21(4) 345–372.

Broekel T, Hartog M, 2013, "Determinants of Cross-Regional R&D Collaboration Networks: An Application of Exponential Random Graph Models", in The Geography of Networks and R&D Collaborations. Ed T Scherngell (Cham, Springer International Publishing) pp 49–70.

Broekel T, Knuepling L, Mewes L, 2023, 'Boosting, sorting and complexity—urban scaling of innovation around the world' Journal of Economic Geography 23(5) 979–1016.

Burger M, van Oort F, Linders G-J, 2009, 'On the Specification of the Gravity Model of Trade: Zeros, Excess Zeros and Zero-inflated Estimation' Spatial Economic Analysis 4(2) 167–190.

Cairncross F, 2001 The death of distance. How the communications revolution is changing our lives (Harvard Business School, Boston, Mass.).

Cantner U, Graf H, 2006, 'The network of innovators in Jena: An application of social network analysis' Research Policy 35(4) 463–480.

Carbonell P, Rodriguez A I, 2006, 'Designing teams for speedy product development: The moderating effect of technological complexity' Journal of Business Research 59(2) 225–232.

Cooke P, Gomez Uranga M, Etxebarria G, 1997, ‘Regional innovation systems: Institutional and organisational dimensions’ Research Policy 26(4-5) 475–491.
Diodato D, Morrison A, Petralia S, 2022, ‘Migration and invention in the Age of Mass Migration’ Journal of Economic Geography 22(2) 477–498.
Esposito C, van der Wouden F, 2022, ‘Learning Amid Technological Disruption: Evidencefrom U.S. Inventors, 1836-1999’ Academy of Management Proceedings 2022(1).
Esposito C R, 2023, ‘The geography of breakthrough invention in the United States over the 20th century’ Research Policy 52(7) 104810.
Fai F, Tunzelmann N von, 2001, ‘Industry-specific competencies and converging technological systems: Evidence from patents’ Structural Change and Economic Dynamics 12(2) 141–170.
Florida R, 2005, ‘The World Is Spiky Globalization has changed the economic playing field, but hasn’t leveled it R Florida Atlantic monthly’ 296(3) 48.
Florida R, Gulden T, Mellander C, 2008, ‘The rise of the mega-region’ Cambridge Journal of Regions, Economy and Society 1(3) 459–476.
Friedman T L, 2005 The world is flat. A brief history of the twenty-first century 1st edn. (Farrar Straus and Giroux, New York, NY).
Gomila R, 2021, ‘Logistic or linear? Estimating causal effects of experimental treatments on binary outcomes using regression analysis’ Journal of experimental psychology. General 150(4) 700–709.
Gordon R J, 2016 The rise and fall of American growth. The U.S. standard of living since the Civil War (Princeton University Press, Princeton and Oxford).
Groot H L F de, Poot J, Smit M J, 2007, ‘Agglomeration, Innovation and Regional Development: Theoretical Perspectives and Meta-Analysis’ SSRN Electronic Journal 38 911.
Hartog M, Gomez A, Orazbayev S, Hausmann R, Crescenzi R, Petralia S, Neal Z, 2022 Inventing modern invention: the professionalization of technological progress in the US (Paper to be presented at DRUID22, Copenhagen Business School, Copenhagen, Denmark).
Hasson J A, Tinbergen J, 1966, ‘Shaping the world economy: suggestions for an international economic policy’ Revue Économique 76(301) 840.
Hidalgo C A, Hausmann R, 2009, ‘The building blocks of economic complexity’ Proceedings of the National Academy of Sciences of the United States of America 106(26) 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.
Jaffe A B, 1989, ‘Real effects of academic research’ American Economic Review 79(5) 957–970.
Jaffe A B, Trajtenberg M, Henderson R, 1993, ‘Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations’ The Quarterly Journal of Economics 108(3) 577–598.
Janssen M J, Abbasiharofteh M, 2022, ‘Boundary spanning R&D collaboration: Key enabling technologies and missions as alleviators of proximity effects?’ Technological Forecasting and Social Change 180(7) 121689.
Lamoreaux N R, Sokoloff K L, 1996, ‘Long-term change in the organization of inventive activity’ Proceedings of the National Academy of Sciences of the United States of America 93(23) 12686–12692.
Li G-C, Lai R, D’Amour A, Doolin D M, Sun Y, Torvik V I, Yu A Z, Fleming L, 2014, ‘Disambiguation and co-authorship networks of the U.S. patent inventor database (1975–2010)’ Research Policy 43(6) 941–955.
Maggioni M A, Nosvelli M, Uberti T E, 2007, ‘Space versus networks in the geography of innovation: A European analysis’ Papers in Regional Science 86(3) 471–493.

Makkonen T, Williams A M, Mitze T, Weidenfeld A, 2018, ‘Science and technology cooperation in cross-border regions: A proximity approach with evidence for Northern Europe’ European Planning Studies 76 1–19.
Melo P C, Graham D J, Noland R B, 2009, ‘A meta-analysis of estimates of urban agglomeration economies’ Regional Science and Urban Economics 39(3) 332–342.
Menzel M-P, Feldman M P, Broekel T, 2017, ‘Institutional change and network evolution: Explorative and exploitative tie formations of co-inventors during the dot-com bubble in the Research Triangle region’ Regional Studies 51(8) 1179–1191.
Mewes L, Broekel T, 2020, ‘Technological complexity and economic growth of regions’ Research Policy 77(5) 104156.
Nathan M, Rosso A, 2015, ‘Mapping digital businesses with big data: Some early findings from the UK’ Research Policy 44(9) 1714–1733.
O’Brien R, 1992 Global financial integration. The end of geography (Council on Foreign Relations Press, New York).
Ohmae K, 1999 The borderless world. Power and strategy in the interlinked economy (HarperBusiness, New York).
Paullin C O, 1932, "Atlas of the Historical Geography of the United States", in . Ed J K Wright (Washington, D.C., Carnegie Institution).
Pavitt K, 1998, ‘Technologies, Products and Organization in the Innovating Firm: What Adam Smith Tells Us and Joseph Schumpeter Doesn't’ Industrial and Corporate Change 7(3) 433–452.
Perry M J, 2014, "The Cost of Air Travel in the US Has Been Remarkably Stable for the Last Decade, and 17% Cheaper Than 20 Years Ago", https://www.aei.org/carpe-diem/the-cost-of-air-travel-in-the-us-has-been-remarkably-stable-for-the-last-decade-and-17-cheaper-than-20-years-ago/.
Petralia S, 2020, ‘Mapping general purpose technologies with patent data’ Research Policy 49(7) 104013.
Petralia S, Balland P-A, Rigby D, 2016a, "HistPat Dataset" (Harvard Dataverse).
Petralia S, Balland P-A, Rigby D L, 2016b, ‘Unveiling the geography of historical patents in the United States from 1836 to 1975’ Scientific data 3 160074.
Prencipe A, 2000, ‘Breadth and depth of technological capabilities in CoPS: The case of the aircraft engine control system’ Research Policy 29(7-8) 895–911.
Rietveld P, Vickerman R, 2003, ‘Transport in regional science: The ?death of distance? is premature’ Papers in Regional Science 83(1) 229–248.
Rodríguez-Pose A, Berlepsch V von, 2015, ‘European Migration, National Origin and Long-term Economic Development in the United States’ Economic Geography 91(4) 393–424.
Scherngell T, Barber M J, 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.
Simensen E O, Abbasiharofteh M, 2022, ‘Sectoral patterns of collaborative tie formation: investigating geographic, cognitive, and technological dimensions’ Industrial and Corporate Change 31(5) 1–36.
Sokoloff K, 1988 Inventive Activity in Early Industrial America. Evidence From Patent Records, 1790 - 1846 (National Bureau of Economic Research, Cambridge, MA).
Sonn J W, Storper M, 2008, ‘The Increasing Importance of Geographical Proximity in Knowledge Production: An Analysis of US Patent Citations, 1975-1997’ Environment and Planning A 40(5) 1020–1039.
Sorenson O, Rivkin J W, Fleming L, 2006, ‘Complexity, networks and knowledge flow’ Research Policy 35(7) 994–1017.
Stanley T D, Jarrell S B, 1989, ‘META-REGRESSION ANALYSIS: A QUANTITATIVE METHOD OF LITERATURE SURVEYS’ Journal of Economic Surveys 3(2) 161–170.

Taylor G R, 1989 The transportation revolution. 1815 - 1860 (Sharpe, Armonk NJ u.a.).
Teaford J C, 1994 Cities of the heartland. The rise and fall of the industrial Midwest 1994th edn. (Indiana University Press, Bloomington Ind.).
Ter Wal A, 2014, 'The dynamics of the inventor network in German biotechnology: Geographic proximity versus triadic closure' Journal of Economic Geography 14(3) 589–620.
Ter Wal A, Boschma R A, 2009, 'Applying social network analysis in economic geography: Framing some key analytic issues' The Annals of Regional Science 43(3) 739–756.
van der Wouden F, 2018 Exploring the Changing Structures of Inventor Collaboration in U.S. Cities between 1836 and 1975 (Los Angeles).
van der Wouden F, 2020, 'A history of collaboration in US invention: Changing patterns of co-invention, complexity and geography' Industrial and Corporate Change 29(3) 599–619.
Viechtbauer W, 2010, 'Conducting Meta-Analyses in R with the metafor Package' Journal of Statistical Software 36(3).
Wuchty S, Jones B F, Uzzi B, 2007, 'The increasing dominance of teams in production of knowledge' Science 316(5827) 1036–1039.

Appendix A. Robustness Checks

To ensure that the evolution of the spatial distribution and structure of inventive activities in the USA do not bias the results of our estimations, we conducted two robustness checks. First, we estimated new models using a subsample of the Northeastern United States, including the states of Maine, New Hampshire, Vermont, Massachusetts, Connecticut, Rhode Island, New York, New Jersey, and Pennsylvania. The regression results align with the original model, suggesting that geographic distance has become an even greater obstacle for interregional collaboration in a smaller U.S. region (see Figure A1 below).

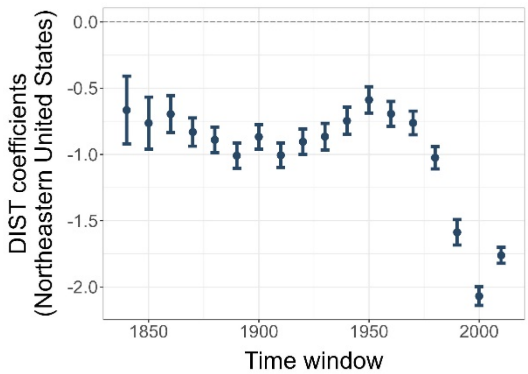

Figure A1. The figure shows the coefficients with the corresponding margin of error for the subsample of the Northeastern United States, including the states of Maine, New Hampshire, Vermont, Massachusetts, Connecticut, Rhode Island, New York, New Jersey, and Pennsylvania.

In the second check, we re-estimated our models, including another control variable, namely, *pot.DIST*. It is operationalized as the mean geographic proximity of a given region to all other regions within the United States in a specific year. The term "geographic distance" is quantified as the linear distance between the centroids of any two regions under consideration. The dynamism inherent in this variable arises from the evolving landscape of regional opportunities for collaborative innovation, which may shift as new regions emerge or as previously inactive regions begin to engage in patenting activities. Additionally, *pot.DIST* is inherently region-specific, reflecting the unique geographic relationality of each region to emerging sites of innovation expansion.

To capture the potential reconfigurations in the spatial concentration of technological activities across the United States over time, we also introduced an augmented variant of *pot.DIST*. In this alternative formulation, we weighted geographic distances by the relative contribution of a region's patent outputs to the aggregate patent landscape of the nation within the same timeframe. Despite the strong correlation ($r > 0.8$) observed between this adjusted measure and the original *pot.DIST* variable, the empirical results remain substantively unchanged (see Figure A2 below). As the figure illustrates, including this control variable does not substantially affect our main findings.

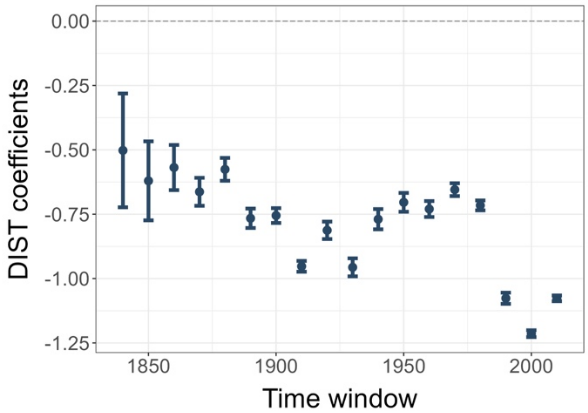

Figure A2. The figure shows the coefficients with the corresponding margin of error for the regression analysis including *pot.DIST* as an additional explanatory variable, i.e., for the case of explicitly controlling for changes in spatial configuration in the USA.