Regional Branching and Local Knowledge Spillovers; a Coevolutionary Process between Relatedness and Co-location, mediated by Complexity

Kevin van Bragt[1]
master's thesis Urban- and Economic Geography
Utrecht University
supervisor: Tom Broekel
words: 8900 (incl. ref)
17 April 2018

Abstract: Despite the burgeoning literature on knowledge spillovers, the regional branching thesis describes the relationship between technological relatedness and co-location as a unidirectional process. However, co-location could equally result in relatedness due to knowledge spillovers. Yet despite emerging insights on 'coevolutionary processes' in the proximity dynamics literature, a dualistic approach to date remains unexplored. Building forth on US patent data from 1850 till 2005 this study will add to the regional diversification literature by examining the 'coevolving', dynamics between two proximity configurations underlying knowledge spillovers and regional branching; technological relatedness and spatial co-location. The results confirm co-location does indeed also influence technological relatedness. Moreover, the results show a significant mediating role of technology age and complexity. Consequently, policy implications and directions for further research include the acknowledgement of regional differences in relatedness, given its 'emergent' characteristics.

keywords: technological proximity, geographical proximity, proximity dynamics, geography of innovation,

[1] J.a.k.vanbragt@students.uu.nl

Introduction

In the past, policymakers have devoted a large amount of attention to the development of 'successful clusters' and high-tech industries. However, many such policies apply a 'one-size-fits-all' approach, and the economic returns to such policies are subject to debate (Crespo et. al., 2017). Storper (1995) argues specific capabilities embedded in the region, referred to as 'untraded interdependencies', are at the basis of region-specific assets and competitiveness. Because regions host a number of spatially bounded, social- and economic processes (e.g. Scott and Storper, 2003; Morgan, 2004), that lead to positive feedback mechanisms (Dobusch and Schüßler, 2013), these untraded interdependencies are influential in determining the future development path of regions due to a path-dependent development (Martin and Sunley, 2006).

Recently, the 'Regional Branching' thesis (Boschma and Frenken, 2009) has made a useful addition to the debate on regional diversification. The regional branching thesis builds forth on this path-dependency by arguing new technological knowledge in a region 'branches out' of the existing technological base of a region towards 'technologically proximate' fields. Following the Proximity School (Torre and Rallet, 2005; Boschma, 2005; Balland et. al., 2014; Broekel, 2014) and knowledge spillover literature (Jaffe, 1986; Jaffe et. al., 1993; Audretsch and Feldman, 1996a), actors search for new knowledge in a geographically-, socially-, organizationally-, institutionally- and cognitively bounded space. Since geographical proximity plays an important role in facilitating, complementing and substituting the other forms of proximities (Boschma and Frenken, 2010; Audretsch and Feldman, 1996a, Boschma, 2005), geographical proximity is an important determinant in the formation of knowledge networks. Moreover, these proximity configurations have been found to be dynamic and interdependent in nature (Balland et. al., 2014). Therefore, some call for the examination of 'coevolutionary proximity dynamics' (Broekel, 2012). After all, "*real evolution, biological or technological, is actually a story of coevolution*" (Kauffman and Macready, 1995, p. 27).

However, despite the acknowledgement of the importance of geographical proximity, the regional branching thesis seems to treat the relationship between technological- and geographical proximity as a uni-directional process. It nevertheless seems highly plausible that geographical proximity could lead to technological proximity. Due to spillover effects and (localized) proximity dynamics, exposure between previously unrelated technologies, and consequently their relatedness, could increase. This process where (local) spillovers lead to relatedness, has to date been largely ignored (yet see Castaldi et. al., 2015), despite the potential existence of a reciprocal, coevolutionary process between both.

In this paper I will therefore add to the debate on regional diversification patterns by empirically examining the extent to which co-location leads to relatedness. I will depart from a point of view where the process of regional branching, and knowledge spillovers, are a mutually dependent processes of emergence and feedback (Levin, 1998). Using the newly available histpat dataset (Petralia et. al., 2016) combined with the NBER patent database (Hall et. al., 2001), I derive measures for technological- and geographical proximity between technologies for over 150 years of innovation in the US. Using panel data regression; long-term, short-term and autocorrelation dynamics are examined.

The results empirically show a reciprocal relationship between technological relatedness and co-location, subsequently pointing out a crucial gap in the proximity dynamics and regional diversification literature. Moreover, the results empirically show a mediating role of technological complexity and technologies' age. The results urge for further research on the interdependent nature of the variables under study. Crucially, the results show relatedness 'emerges' out of co-location, which points towards future research directions that could increase knowledge on regional technological diversification, and the application of region specific relatedness to determine opportunities.

The next section will present the relevant literature. The third section will outline the data and methodological approach. The fourth section outlines and interprets the results. The last section will discuss the validity of the current study, and discuss the implications for further research and policy.

Theory

Technological Relatedness & Regional Branching

Innovation is frequently compared to the biological process of evolution. Wagner and Rosen go as far as to argue that '*the process of innovation reflects almost everything we have learned about biological evolution*' (2014, p. 2). One of the most typical similarities is the view that innovation results from a (re-)combination of previously unrelated technologies (Nelson and Winter, 1982). This recombination resembles that of bacteria, taking place horizontally between many different individuals (Wagner and Rosen, 2014). The already existing pool of technologies, like genes, forms the 'building blocks' for future recombination efforts (Strumsky et. al., 2015), upon which development of new technologies is dependent for recombination (Nelson and Winter, 1982).

However, not all technologies combine in the same manner, nor with equal or random probability. In their seminal work, Engelsman and van Raan (1994) exploit the classification of patents in technological classes indicating underlying knowledge, to map the technological relatedness between these classes in what is referred to as the knowledge- or technology space (e.g. Kogler et. al., 2013). The technology space is a network representation of the proximity between technological classes, referred to as relatedness. Within this network, technologies can be related due to a variety of reasons. In a study on firm diversification, Breschi et. al. (2003) mention three[2] potential reasons for technological proximity. To start, 'knowledge commonalities' imply a process where firms' product diversification tends to steer towards products that require similar capabilities, problem solving, or heuristics of search. In this case, the same piece of knowledge might be applied in different new inventions, resulting in economies of scope. Second, knowledge complementarities imply a relatively high recombination potential between technologies. In this instance, the 'whole is more than the sum of it's part'. Third, knowledge proximity refers to both intended- and unintended learning outcomes. Unintended learning occurs through knowledge spillovers resulting from firms' innovative activities, intended learning results from search processes, often in fields that are somehow related to their current economic activities.

[2] Empirically distinguishing between these different forms of relatedness could prove a meaningful addition to the current effort, unfortunately, no method to empirically do so has been put forward yet.

Using countries' export data, Hidalgo et. al. (2007) combine mentioned mapping of proximities with the revealed comparative advantage (RCA) measure to construct the 'product space', which identifies proximities between different export products. Their findings suggest that countries tend to diversify towards products that are 'proximate' to their current current export basket. As with technological relatedness, the exact reason for given proximities remains uncertain[3]. Regions have been shown to follow a similar path-dependent development progress where current capabilities influence the opportunities for, and direction of, future diversification. Within the framework of evolutionary economic geography, technological relatedness has therefore achieved an important role in describing why certain locations develop a specific sort of knowledge (e.g. Kogler et. al., 2017; Neffke et. al., 2011, Boschma, 2017). Boschma and Frenken (2009) refer to this process of regional related diversification as 'regional branching'. Boschma et. al. (2014) empirically show that when relatedness increases, probability of entry increases and exit probabilities decrease. Four different mechanisms through which technological relatedness influences regional diversification patterns have been identified. Although I will briefly discuss each, for a more detailed overview I would like to refer to Boschma & Gianelle (2014). To start, firm diversification leads firms to diversify towards related fields as a result of learning processes (Breschi et. al., 2003). Secondly, spin-offs tend to build on rather similar though slightly different knowledge, spin-offs therefore tend to start in related industries. Thirdly, labour mobility is a crucial mechanism for transferring knowledge, skills and experience between firms and industries within regions. Lastly, social- and collaboration networks may lead to social- and cognitive proximity on an actor level, hence increasing collaboration (Boschma and Gianelle, 2014).

Local Knowledge Spillovers and Dynamic Proximity Configurations

The importance of geographical proximity for learning and innovation networks has long been recognized. Marshall (1890) claimed "*The mysteries of the trade become no mysteries; but are as it were in the air"* (chapter 10). Arrow (1962 in: Glaeser et. al., 1992) ascribes this observation as a result of the non-exclusive and non-rival nature of knowledge. However, in explaining knowledge spillovers both emphasized intra-industry spillovers. Only after the pioneering work of Jacobs (1969), the role of diversity and inter-industry spillovers became more apparent. Jacobs additionally ascribes localized processes of learning to the difference between information and knowledge, the latter of which is embedded in individuals, firms and organizational routines, and therefore less easily transferred. Transmission of such embedded, or 'sticky' (von Hipple, 1994), knowledge can be enhanced by face-2-face contact (Audretsch and Feldman, 1996a). Hence, knowledge does not flow freely out of its own, rather it needs to be embedded and driven by actors.

The Proximity Literature (e.g. Torre and Rallet, 2005; Boschma, 2005) builds forth on these findings by arguing geographical proximity does not only have a direct-, but also an indirect effect on learning and knowledge networks. The indirect effect of geographical proximity is due to the interdependencies that exist between different forms of proximities. Boschma (2005) mentions four other forms of proximity; organizational, institutional, social

[3] Hidalgo et. al. (2007) mention as possible reasons; 'physical capital, labor, land, skills or human capital, infrastructure, and institutions'

and cognitive[4], which also influence the formation of knowledge networks (Boschma and Frenken, 2010). For these proximity configurations it is argued that: "*too much and too little proximity are both detrimental to learning and innovation*" (Boschma, 2005, p. 71). However, the different forms of proximities can compensate for one another. Geographical proximity plays an especially important role in this compensating, since it can facilitate, complement and substitute other types of proximities (Broekel, 2012). Empirically, geographical proximity has been shown to influence social- (Morgan, 2004) and cognitive proximity (Audretsch and feldman, 1996a), and be interdependent with organizational- and institutional proximity (Torre and Rallet, 2005). Orlando (2004) adds technological proximity as a sixth proximity configuration to the list of those that interact with geographical proximity. Interestingly, Orlando finds technological proximity can act as a substitute for geographical proximity. Cassi and Plunket (2015) refer to interactions within proximities as 'closure' and refer to 'bridging' as interactions between different types of proximities.

Balland et. al. (2014) add to the literature by questioning the 'uni-causal' logic underlying much of the earlier proximity literature, which assumes (dynamic) knowledge networks result from (static) proximity configurations. Rather, Balland et. al. argue, proximity configurations themselves are subject to change as well due to the constant reconfigurations of knowledge networks, and are consequently 'dynamic' in nature. These dynamics, explain the interdependencies between different forms of proximity, as they imply that changes in a particular proximity configuration, through changing the knowledge networks, could lead to consecutive changes in other proximity configurations. Broekel (2015) describes these reciprocally interdependent proximities as 'coevolving', and describes three types; simultaneous coevolution implies a correlation between changes in a single time period, long term coevolution describes reciprocal changes occuring in a subsequent period, temporal autocorrelation dynamics imply that if a configuration changes in a certain period, this same configuration is likely to change in the subsequent period as well.

Technological Relatedness as an Emergent Property

Within the Complex Adaptive Systems framework, Levin (1998) raises awareness for the relationship between structure and functioning. "*Macroscopic system properties … emerge from interactions among components, and may feed-back to influence the subsequent development of those interactions*" (p. 431). While the literature on region branching seems to assume relatedness influences co-location in a unicausal manner, the process described by Levin seems to more accurately describe their causal relationship. Different 'bits' of technological knowledge make up the diversity among which 'local' interactions take place. These interactions between technologies are of course embedded in the the actors that carry knowledge. Hence, the five different forms of proximity have a large influence on which interactions take place. In other words, both directly and indirectly, *geographical proximity to a large extent influences what interactions take place between technologies*. Technologies that are highly co-located, have a larger chance to be recombined due to spillover effects following their proximity configurations. This was already reflected in the work of Jacobs (1969), who argues localized diversity enables the (re-)combination, or 'cross-fertilization', of different knowledge and ideas. After going through a process of trial and error, successful

[4] It is beyond the scope of the present paper to go into the different types of proximities and their linkages in detail. For a concise overview I refer to Boschma (2005) and Boschma and Frenken (2010).

recombinations lead to new technological breakthroughs and applications that, if adopted, influence the ‘global’ measure of relatedness (e.g. the pattern of the technology space). This in turn feeds back through the branching processes as described in the literature. Hypothesis one will test this proposition.

Hypothesis 1: “spillover effects and regional branching form a positive coevolutionary feedback loop”

The effects of Technologies’ Life-Cycle

Some argue the need for geographical proximity is dependent on the type of innovation (Torre and Rallet, 2005; Castaldi et. al., 2015) and related to that, the age or life-cycle phase technologies are at (Ter Wal, 2013; Puga, 2001; Audretsch and Feldman, 1996b). Puga (2001) introduces the concept of ‘nursery cities’. Nursery cities are highly diverse cities which although less cost efficient, offer new firms the possibility to ‘invest’ in learning and exploration in their early development stages. Castaldi et. al. (2015) find that unrelated variety increases the likelihood of technological breakthroughs. These findings are in line with the idea of Jacobs (1969) that diversity breeds recombination potential. Audretsch and Feldman (1996b) find that industries tend to concentrate less in space as they mature. Ter Wal (2013) argues geographical proximity is crucial especially in the early stages of an industry, while triadic closure (co-operating with partners’ partners) becomes more prominent in later stages. Torre and Rallet (2005) argue that the role of face-to-face contact for knowledge transmission depends on the type of innovation process, with in decreasing order; exploration, exploitation and imitation. These earlier findings all point towards the fact that geographical proximity is crucial especially to establish technological compatibility. However after technologies have successfully emerged, the emphasis might shift towards other forms of proximity. As Menzel (2008) points out, temporary clusters such as conferences or industry fairs can serve as a means to provide temporary proximity (Torre and Rallet, 2005). Building forth on the theory, we can thus expect new technologies to be creating new relatedness (patterns), often in a short time period (Castaldi et. al., 2015; Schumpeter, 1942), resulting from spillovers (Jacobs, 1969) and therefore initiated by geographical proximity. However, as technologies age and the type of innovation changes, the need for geographical proximity might become a less crucial factor, while patterns in the technology space might become a more important driver of regional technological diversification.

Hypothesis 2: “When Technologies age this results in a shift in emphasis from co-location towards relatedness”

Technological Complexity as Mediating Factor

Sonn and Storper (2008) show that despite the improvements in ICT-technologies, the proportion of local citations has increased rather than decreased between 1975 and 1997. Some argue technological complexity increases the difficulty of passive learning (Pintea and Thompson, 2007), making complex knowledge more ‘sticky’ (Balland and Rigby, 2017). Recent studies refer to this technological complexity and attempt to measure technologies’ complexity (e.g. Kauffman and Macready, 1995; Fleming and Sorenson, 2001; Balland and

Rigby, 2017; Broekel, 2017), as well as the influence of technological complexity on a number of economic processes (e.g. Pintea and Thompson, 2007).

Kauffman and Macready (1995) compare technological innovation with biological evolution, and describe innovation as a search for ‘peaks’ within the ‘fitness landscape’. The topology of this fitness landscape is dependent upon the amount of components and the interdependency between these components. With an increasing amount of components making up this fitness landscape, the amount of peaks greatly increases, while the average height of these peaks declines (Kauffman, 1993). When interdependence increases, the ‘whole is more than the sum of its parts’, and differences in the heights of the peaks increase. Consequently, local search is favorable for less interdependent landscapes, as the landscape contains less variation in the height of its peaks, and hence incremental innovation is preferred. However, as interdependence increases, the amount of peaks increases while only some are relatively high. Consequently, with high interdependence, local search will be less useful (Fleming and Sorenson, 2001), as it could lead to (suboptimal) technological lock-ins. This implies that higher interdependency could result in a need to bridge ‘distances’.

Broekel (2017) argues complexity increases over time, and that complex technologies require more R&D and more cooperation. Complex technologies require more R&D because the difficulty of ‘incremental improvement’, that is improvement on the same ‘hill’ in the fitness landscape, doubles with each incremental improvement (Kauffman and Macready, 1995). More cooperation is needed because the amount of different bits of knowledge required increases (Broekel, 2017). Given bounded rationality, because R&D and innovation are subject to increasing returns to scale (Schumpeter, 1942), and because geographical proximity serves as a strong facilitator for learning and cooperation and thus spatially constraints diffusion (Balland and Rigby, 2017), it makes sense that complex technologies concentrate in space (Broekel, 2017). Balland and Rigby (2017) find empirical evidence that complex technologies are more ‘sticky’ and concentrate in space, implying complexity thus constraints diffusion. They show that citing between patents in different localities is less likely when it involves complex knowledge, hence complex technologies have been found to be less mobile than non-complex technologies. These findings could be interpreted as a result from the tacitness of knowledge (von Hipple, 1994), because of which the need for face-2-face contact is higher. Hence, for complex technologies, it can be expected that co-location is extra important.

Following the idea of the fitness landscape, in order for a locality to diversify towards more complex technologies, technologies will have to be (re)combined in ‘better’ configurations. As pointed out, interdependence leads to many possible sub-optimal peaks. As higher complexity implies more different bits of knowledge, distances somehow have to be bridged. It seems likely that even specialized localities therefore need some ‘influx’ of new knowledge. However, it seems likely that with growing complexity, a relatively higher compatibility is needed to obtain a successful improvement. Since relatedness is a way to approach this compatibility, I expect an increase in the influence of relatedness on co-location with increasing complexity.

Hypothesis 3: “Technological Complexity amplifies the coevolutionary process between relatedness and co-location”

Empirics

Data and Variables

The difficulty of empirically measuring knowledge flows and networks has long been recognized (e.g. Jaffe, 1986). However, patent data has emerged as an accepted way of doing so (Griliches, 1990). For this study, I will therefore make use of the Histpat dataset (Petralia et. al., 2016) combined with patent data provided by the United States Patent and Trademark Office [USPTO]; the CPC classification file (USPTO, 2018) and the NBER patent data (Hall et. al., 2001). Consequently, data on the location[5] and classification (CPC) of patents granted in the US from 1840 till 2005 is constructed. Patent classifications can be considered as indicating different 'bits' of knowledge, and consequently they can be used to obtain information about the specific type of technological knowledge on which a patent builds (see e.g. Jaffe, 1986). Since the used data also includes geographical location, it is possible to subsequently examine the occurrence of classes within locations. Given this study's focus on regional (diversification of) technological capabilities, this is crucial information. By building forth on network analysis methods, proximity configurations among the different classes can be calculated in both the technological- and the geographical dimension.

Since these proximity dimensions are dyadic in their nature, the main units of analysis are not those individual CPC classes, but rather it is every possible dyad between these individual classes, amounting up to n^2-n pairs of the n CPC classes under study. For most of the technological classes under study, especially in earlier time periods, the amount of patents assigned to each specific class are either (i) relatively small, (ii) or the patent class does not yet exist at all. Because the calculated variables are relative measures, small differences can result in a large variability. Therefore a tradeoff has to be made between aggregating and thus losing detail in the data on the one hand, while on the other hand allowing a too fine grained level of technological knowledge will result in a large variability. I therefore make use of the main, 4-digit class level (n=641). This 4-digit level is considered to be detailed enough to capture different types of knowledge that require cooperation, while at the same time restricting to a certain extent the amount of zero's appearing in the data. In addition, variables are calculated over five year periods in order to get more stable values.

Technological Relatedness

Following the pioneering work of Engelsman and van Raan (1994), this study will make use of a co-occurrence analysis of patent classifications within patents. Since the total amount of (co-)occurrences of technological classes are highly heterogeneous, the co-occurrences (C_{ij}) have to be normalized to account for this size effect (S) of the classes i and j involved in the dyad, as well as for the total amount of co-occurrences (T) in the data. In order to get any meaningful values for the link strengths. Van Eck and Waltman (2009) find that for normalizing co-occurrence counts, probabilistic similarity measures are preferred. Equation one depicts the probabilistic method for normalization used in this study[6].

[5] For the data prior to 1977, the locations of patents could be either applicant or inventor. For data after 1977 locations are those of the inventor.
[6] Following the method used in the EconGeo Package for R (Balland, 2017).

(1) $$RELij = \frac{Cij}{(Si/T)\cdot(Sj/T) + (Sj/T)\cdot(Si/T - Sj)\cdot(T/2)}$$

Co-location

Given the emphasis on processes that are highly dependent on labour mobility and commuting ties, the geographical unit of analysis will be the 'core based statistical areas' (cbsa) as provided by the US census bureau. Cbsa's consist of a group of US counties that share at least one core area with at least 10.000 inhabitants. More importantly, adjacent cbsa counties share important social- and economic linkages, measured in commuting ties (United States Census Bureau, 2018). Since patents can be ascribed both multiple locations and multiple classes, weights are assigned to the occurrence of classes within regions in such a way that they add up to unity for each individual patent[7].

To measure co-location of technologies in a region, I will combine the RTA measure with a co-occurrence analysis as described above. The RTA measures the share of technological output that is ascribed to class *i* in the region's r total patent output, divided by the share this technology class has in the total patent output of the complete sample. In other words, the RTA reflects whether the share of technology i in region r is larger than the share of i 'globally'. When RTA < 1 this implies a technological disadvantage of region r in technology class *i*, if RTA > 1 this implies a technological advantage. A relatively high RTA implies that a region possesses specific assets that enable this relatively high patenting rate. Following Balland and Rigby (2017) The RTA of region *r* in patent class *i* is given by equation two. Here P is the amount of patents, *i* is the technological class, and *r* is the region.

(2) $$RTA\ r,i = \frac{P r,i / \Sigma i\ P r,i}{\Sigma r\ P r,i / \Sigma r\ \Sigma i\ P r,i}$$

To derive the final measure geographical co-location between *i* and *j*, a co-occurrence analysis (see eq. 1) will be performed to measure to what extent technologies tend to be 'co-located' within regions. Only technological classes for which a region has a revealed technological advantage (RTA > 1) will be considered. Note that such a measure treats geographical proximity as a duality rather than a measure of decaying geographic distance. It empirically reflects the tendency for two classes to be overrepresented within the same regions. Also note that co-location does not imply concentration, as the measure is dyadic, nor co-concentration, given the fact that high RTA values are set to one.

Technological Complexity

Different measures of technological complexity exist. Fleming and Sorenson (2001) argue technologies are more complex if recombination is difficult to achieve. Consequently, Fleming and Sorenson assume that complex recombinations are more scarce than simple ones. Based on the 'fitness landscape model' (see e.g. Kauffman and MacReady, 1995), they derive a measure that approaches this hypothetical 'ease of recombination' of individual patent classes based on past recombinations. Following Hidalgo and Hausmann (2009), Balland and Rigby (2017) put forward a 'spatial approach' to measuring complexity. The

[7] For example, a patent with three classes and two locations gives 0.167 times each class for each location.

theoretical argument is that producing a technology requires ‘building blocks’, some of these blocks or their combinations are more complex than others. To produce complex technologies, the constraints due to the required building blocks are larger, and therefore less regions produce complex technologies. At the same time, the diversity of a region is assumed to represent the total amount of capabilities (building blocks) possessed by that region. Balland and Rigby use the RTA’s of regions to calculate ubiquity of technologies and diversity of regions. By iteratively ‘reflecting’ the ubiquity and diversity, they derive a complexity measure. Hence, rather than scarce in frequency, Balland and Rigby assume spatial scarcity of complex technologies.

Broekel (2017) criticises both methods for the fundamental role of scarcity in the measurements. He argues less frequent combinations might also be the result of little economic or technological interest. Technologies’ spatial distribution is driven by other factors as well, such as history, policy and geography. Furthermore, measuring complexity based on spatial scarcity results in high endogeneity when analysing spatial phenomena, making it impossible to test the assumption of complexity truly being spatially scarce. Broekel (2017) therefore proposes a third alternative, referred to as structural complexity. Similar to Fleming and Sorenson (2001), Broekel assumes innovation is a process of recombination, where complexity is determined by the difficulty of recombination. However, rather than using past recombinations, recombination difficulty is proxied by measuring the structural complexity of the recombination network among technological classes. For each technological class c respectively, a binary matrix G_c is constructed containing the co-occurrences between subclasses for the subset of patents containing class c specifically. G_c therefore represents the ‘combinatorial network’ of technology c. In order to measure the complexity of G_c a composite of four network variables, the Network Diversity Score [NDS], is calculated (Emmert-Streib and Dehmer, 2012 in: Broekel, 2017). To increase robustness, this is done for a series of samples from G_c.

The structural complexity measure has multiple advantages over the other two measures for application in this study. Firstly, the structural complexity measure circumvents the potential bias introduced by the scarcity assumption. Especially when combined with the relatedness variable based on patent co-occurrences (recombination frequency) and co-location on RTA’s (spatial frequency), this is an important determinant. Secondly, variability of the structural measure is considerably lower than that of its peers (Broekel, 2017). The complexity values calculated in Broekel (2017b) for CPC classes from 1840 till 2010 have therefore been used in the current study.

Life-Cycle position

In order to approach technologies’ life-cycle position, the average technological age of each dyadic pair *ij* is calculated. For each specific technological class *i*, age is calculated by taking the average age of all patents that have been assigned this specific class *i*. The variable AGE_{ij} then represents the average age of *i* and *j* combined.

Broekel et. al. (2014) outline four methods[8] that are commonly used in economic geography to examine knowledge and innovation networks. Especially promising for modelling network (co)evolution is the SAOM model (see e.g. Snijders, 2014). The SAOM model allows to combine variables from different network levels, that is; actor specific characteristics, dyadic, and network or graph level statistics. However, since technological classes are subject to random 'evolution', they are not real actors, and such a model would build on unrealistic assumptions. Another commonly used methodology is the gravity model, currently well established in economics to explain the quantity of trade between countries. Moreover, studies on migration have similarly used the gravity model to explain the quantity of migration in relation to trade (e.g Fagiolo and Mastrorillo, 2014)). Hence, gravity models are currently used to describe proximity (dynamics) between different 'networks'. More recent applications have shifted from focussing on locational attributes towards trying to include measures inspired by network analysis (e.g. Dueñas and Fagiolo, 2011), such as for instance triadic closure (Ter Wal, 2009). Nevertheless, it is well known that modelling dyadic data brings with it difficulties that result from violating the i.i.d. assumption (Wasserman and Faust, 1994), as errors are highly interdependent. As it turns out the error terms are highly heteroskedastic when examining using the gravity model[9] (using different distribution families). Consequently an alternative to the common approaches has been employed.

The longitudinal nature of the data has three main advantages. Firstly, it allows to regress rates of change rather than absolute values, which is preferred given the dynamic nature of the processes under study. After all, a certain proximity relation might be dependent on other characteristics, such as for instance AGE. A 'static' approach of measuring correlation between the variables would give an (unrealistically) persistent view of these influences/interdependencies. Moreover, using growth rates eases interpretation despite the large variability present in the observed data. Secondly, (absolute) starting values can be included in order to account for path-dependency. Thirdly, and crucially, it can serve as a way to control for both time- and individual unobserved heterogeneity present due to omitted variable bias. Large variability between technologies' characteristics cannot be measured, nor are time-related trends related to economic, political and technological events included, such as i.a. the emergence of new technological paradigms and GPT's.

Consequently, a dynamic panel data approach is employed. In order to account for both time and individual unobserved heterogeneity, a two-way fixed effects model is employed where λt and αi are the time and individual fixed effects dummies[10]. While alternative methods[11] would have been the within and first difference estimators, this would mean a loss of information on the time-fixed effects. With relatedness and co-location being X and Y depending on the direction, the regression equation can be presented in a 'bidirectional way' as given in equation 3. Regardless of the dependent variable [DV]; $\Delta X_{ij,t}$

[8] These are the; Gravity Model, MRQAP regression, ERGM's and SAOM's, see Broekel et. al. (2014) for a detailed overview.

[9] This implies an omitted variable bias, which given the interdependent data, might be solved using network measures that have been found common in networks (evolution).

[10] For OLS regression, the hausman test rejects the null hypothesis of random effects with $p < 2.2*10^{16}$.

[11] Moreover, as the 'population' of technological classes under study is dynamic (after all, classes emerge and disappear over time), only pairs that are non-zero over three consecutive periods are considered. The panel is therefore highly unbalanced and fixed effects dummies are preferred.

represents simultaneous coevolution, $\Delta X_{ij,t-1}$ long term coevolution, $\Delta Y_{ij,t-1}$ is used to examine temporal autocorrelation, $\Delta X_{ij,t-2}$ and $\Delta Y_{ij,t-2}$ represent the starting values. The starting values make sense when assuming path-dependency, but more importantly, combined with growth rates combined, lagged (dependent) variables serve as instrumental variables reducing potential endogeneity bias due to the simultaneity of the variables. Note that for growth relatedness as DV, the process of knowledge spillovers is approximated, while for growth co-location as DV regional branching is considered. Also note technological complexity is not included. From the theory it is expected that complexity influences not merely the response variable, but instead alters the causal relationship between relatedness and co-location. To give insight in this mediating role of technological complexity, interaction effects between a dummy variable for complexity and the independent variables relating to co-location and relatedness are added[12]. The dummy variable for complexity is set to 1 if the structural complexity $COM_{ij,t} > COM_t$.

(3) $\Delta Y_{ij,t} = \alpha_i + \lambda_t + \beta_1 \Delta Y_{ij,t-1} + \beta_2 \Delta X_{ij,t-1} + \beta_3 \Delta X_{ij,t} + \beta_4 Y_{ij,t-1} + \beta_5 X_{ij,t-1} + \beta_6 AGE_{ij,t} + \mu_{ij,t}$

As it turns out the error terms are highly heteroskedastic when using OLS[13], a quantile regression is employed as method of estimation (see Koenker and Basset, 1978). While OLS regression is based on the squared deviation from the conditional mean, quantile regression is based on the absolute deviation. Quantile regression allows the estimation of coefficients for different quantiles of the response variables separately[14]. As a result, quantile regression is less sensitive to outliers, Consequently, quantile regression is less sensitive to outliers (Koenker and Basset, 1978), and robust to heteroskedasticity (Koenker and Hallock, 2001). Koenker (2004) outlines the application of quantile regression to panel data with fixed effects[15].

Results & Discussion

The Coevolutionary Process between Co-location and Relatedness

Table one (next page) shows all estimated coefficients, except for AGE and $\Delta Relatedness_{ij,t}$ on growth co-location are significant on the 99% level. For both relatedness growth and co-location growth as DV, the beta's for $\Delta Y_{ij,t-1}$ on ΔY are negative, which implies a process of negative temporal autocorrelation. In contrast, positive beta's are observed for $\Delta X_{ij,t-1}$ and $\Delta X_{ij,t}$ on ΔY in both cases. Hence, positive feedback exists between relatedness and co-location in both directions and both simultaneous and long-term. Negative temporal autocorrelation serves a stabilizing/suppressing role. The relatively large magnitudes of the beta's for long-term influence compared to short-term imply this process is mostly a long-term one. Although this seems to be more evident for the influence of growth relatedness on growth co-location than that of growth co-location on growth relatedness. Beta's for the starting values are positive for $X_{ij,t-2}$ and negative for $Y_{ij,t-2}$.

---

[12] Alternative compositions can be found in table two in the appendix.
[13] Arguably because of simultaneity bias
[14] Hence, the regression coefficient for explanatory variable X is given for a specific quantile of the response variable Y. Different quantiles could be examined to test e.g. for U-shaped patterns. However, this is beyond the scope of the present paper.
[15] The Quantreg package (Koenker, 2018) has been used for estimation.

Simply put, these results confirm the existence of a coevolutionary process between relatedness and co-location. Not only does relatedness influence regional diversification through a branching process, locally available capabilities also lead to recombination of previously unrelated technologies. Hence, the concept of emergent macro-properties as a result of local interactions, as outlined by Levin (1998), seems to apply to co-localization of technologies and technological relatedness as well. However, since growth is negatively autocorrelated, this effect is somewhat suppressed in the long run, matching the idea that neither too much nor too little proximity is beneficial for innovation (Boschma, 2005). This idea is further confirmed by the negative influence of the starting values for $Y_{ij,t-2}$. Again, the starting values for $X_{ij,t-2}$ have a positive influence, showing a long-term tendency to positive feedback mechanisms.

Table 1: Panel Model Regression Coefficients[16]

	Dependent variable:	
	$\Delta Relatedness_{ij,t}$	$\Delta Co-location_{ij,t}$
$AGE(t)$	.328***	-.076***
$\Delta Relatedness_{ij,t}$	-	.007***
$\Delta Co-location_{ij,t}$	.126***	-
$\Delta Relatedness_{ij,t-1}$	-.126***	.003***
$\Delta Co-location_{ij,t-1}$	.065***	-.434***
$Co-location_{ij,t-2}$	3.297***	-15.739***
$Relatedness_{ij,t-2}$	-.028***	.030***
$COMij * \Delta Relatedness_{ij,t}$	-	-.002
$COMij * \Delta Co-location_{ij,t}$	-.088***	-
$COMij * \Delta Relatedness_{ij,t-1}$	.024***	.001
$COMij * \Delta Co-location_{ij,t-1}$	-.064***	-.022***
$COMij * Relatedness_{ij,t-2}$	-.007*	.061***
$COMij * Co-location_{ij,t-2}$	-5.688***	-2.616***

Note: 5 year lags, tau = 0.5, p<0.1, p<0.05**, p<0.01***, time- and individual fixed effects included, n=236.916*

The Influence of Age on Co-location and Technological Proximity

For AGE, a positive influence is observed on relatedness growth, while a negative influence is observed for co-location growth. This matches the expectations derived from the theory, since ‘older technologies concentrate less in space’ (Audretsch and Feldman, 1996b). It seems highly plausible that new technologies develop in nursery cities, as described by Duranton and Puga (2001). Subsequently, if technologies develop, they establish themselves within the technology space, influenced by the technologies they are co-located with. The type of innovation switches from exploration to exploitation to imitation (Torre and

[16] In order to ease interpretation despite high variance, variables are mean centered. time dummies for growth relatedness range between 15 and 60, time dummies for growth colocation range between -25 and -40.

Rallet, 2005). Or as Castaldi et. al. (2015) put it, breakthrough technologies appear due to recombination of unrelated technologies, and might become related over time. These results also match those of Ter Wal (2013) and Orlando (2004). One might also interpret these results with the view that this initial phase of radical innovation is a search for a 'new' peak in the fitness landscape, once a position on this landscape is taken, the type of innovation might change towards incremental innovation that builds forth on the initially established patterns. However, in line with Torre and Rallet (2005), this might be a consequence of a shift towards temporary proximity. Temporary proximities could equally result in an emergent pattern of relatedness.

The Mediating Role of Complexity

The interaction effects with complexity show that for both DV's the effect of age is amplified for complex technologies, the effect on co-location is not significant however. Simultaneous coevolution is, in both cases, less for complex technologies. Long term coevolution is to a very large extent suppressed for relatedness growth as DV. For growth co-location as DV long-term co-evolution is increased by complexity, but only significant on the long-term. For growth relatedness, complexity decreases the extent of negative temporal autocorrelation, while for growth co-location complexity increases negative temporal autocorrelation. Lastly, for co-location as DV complexity amplifies the effect of both starting values. For growth relatedness as DV the negative effect of $Y_{ij,t-2}$ is slightly decreased while the effect of $X_{ij,t-2}$ becomes negative for complex dyads.

For the process regional branching, only the effect of the starting value $X_{ij,t-2}$ is significant. Hence, the branching process seems is amplified by complexity, while the emphasis seems to shift towards the long-term. This matches the expectations that face-to-face contact is more important for complex technologies. As learning and cooperation is improved by co-location, actors might to a larger extent actively locate in an area devoted to a specific technological domain when dealing with complex technologies. This explanation is in line with the finding that complex knowledge is 'sticky' (Balland and Rigby, 2017).

The effect of co-location on relatedness, both short- and long-term, is greatly reduced for complex technologies. This contrasts the expectation that complexity would increase the need for face-2-face contact and hence increase the causal relationship. The suppressing effect technological complexity has on the effect of co-location on relatedness growth could be explained by the inverted U-shaped pattern that exists between returns to spatial proximity and complexity (Sorenson et. al., 2006). Perhaps a mix between local and non-local knowledge (Bathelt et. al., 2004) better suits complex technologies, as paths for potential development are less in quantity (Kaufmann and MacReady, 1995), and the likelihood of (suboptimal) lock-in effects is thus larger. Since complex technologies require more collaboration (Broekel, 2017), and the integration of more 'different' types of knowledge, actors might benefit from a higher degree of 'bridging' to fill knowledge gaps (Cassi and Plunket, 2015). Following the view of the fitness landscape, this would imply making a jump towards a different 'peak' (i.e. leading to a breakthrough development). After all, the main driver of changes in proximity configurations is bridging (Menzel, 2008), and proximity also has a temporary dimension (Torre and Rallet, 2005). Face-2-face contact could therefore also be stimulated by organized events such as conferences (Menzel, 2008). This would allow actors to overcome negative lock-in effects.

Conclusion

This study has empirically shown that co-location of technologies can lead to technological relatedness, a process often overlooked in the scientific literature. Moreover, both age and complexity have been found to have significant effects on the causal mechanisms between relatedness and co-location. Combining the current results with previous literature suggests co-location plays an important role in the development of new (non-complex) technologies. However, as complexity increases, bridging knowledge gaps in different ways might become more important in order to avoid suboptimal ("local") lock-in on the fitness landscape. In the words of Levin (1998), this coevolutionary process results in the emergent property of relatedness, which consequently feeds back to local interactions.

However, the current study is not without its limitations. As discussed by Griliches (1990) Hall et. al. (2000) and Pavitt (1988), the use of patent data is subject to debate. An important disadvantage of using patent data for the current study is that firms tend to 'split' their patents among locations, this could potentially introduce noise in the data used for location. Moreover, patenting behaviour is subject to change over time (Pavitt, 1988). Apart from these data-related issues, future research could improve on this study in a number of ways; To start, while arguably an important determinant in innovation networks, this study fails to account for 'temporary proximity' (e.g. Torre and Rallet, 2005). Secondly, while this study builds on the use of network data, the application of network analysis methods is limited in scope. Further research could make use of recent developments in the field to empirically examine the (network) processes occurring. Moreover, this might be combined for a multitude of different networks representing the different forms of proximity. Thirdly, the methodology of this study has build on highly skewed ratio variables, which constrained the possible methods for examination. Nevertheless, using logit methods, such difficulties might be overcome and more advanced methods could be applied. Fourth, it can reasonably be expected that the examined processes take place between more generally defined industries, using groups of patent classes to indicate industries (e.g. SIC's) could give further insight[17], as well as provide more practical knowledge for policy decisions. Lastly, it is not examined whether significant differences in patent value or originality (available via NBER) exist as a consequence of the examined proximity configurations. Such an approach could give additional insights in 'breakthrough patents', and further confirm results from this and other studies.

Despite its shortcomings, this study has shown an important gap in the current scientific literature that asks for further attention. Relatedness is considered to influence regional co-location in a 'top-down' manner. However, by comparing the relationship between relatedness and co-location to a system of feedback and relatedness as an emergent property, it can be expected that much like the different development paths taken by species depending on their location/environment, relatedness is not the same at every locality. Biologists refer to a process of 'adaptive speciation', when a single species mutates

[17] This would have the additional benefit of solving to a large extent the amount of zero's measured and issues with outliers.

into different directions, evolves towards different fitness peaks, and ultimately different species (e.g. Weissing et. al., 2011). Ofcourse, this idea is not new to Economic Geographers, as path-dependency plays an important role in describing spatial development paths. However, while seemingly taken for granted, the 'global' measure of relatedness might be suboptimal, as this implies differences in locational opportunities. An interesting direction for further research would be the examination of a more 'local' measure of relatedness. Building forth on the extensive literature and methodology developed by (evolutionary) biologists and ecologists, this could be compared to (i) the 'global' technology space, (ii) regional diversification, and (iii) local measures of relatedness among each other. Insights might be relevant for those interested in; radical- and incremental innovation, the diversification vs. specialization debate and technological lock-in effects.

Bibliography

Arellano, M. & S. Bonhomme (2017). Quantile Selection Models with an Application to understanding Changes in Wage Inequality. Econometrica 81(1), pp. 1-28.

Audretsch, D. & M.P. Feldman (1996a). R&D Spillovers and the Geography of Innovation and Production. American Economic Review, 86(3), pp. 630-40.

Audretsch, D & M.P. Feldman (1996b). Innovative Clusters and the Industry Life Cycle. Review of Industrial Organization 11(2), pp. 253-273.

Balland, P.A. (2017). Economic Geography in R: Introduction to the EconGeo package. Papers in Evolutionary Economic Geography #17.09, pp. 1-75.

Balland, P.A., R. Boschma, J. Crespo, & D. Rigby (2017). Smart Specialization Policy in the European Union: Relatedness, Knowledge complexity and Regional Diversification. Regional Studies, *forthcoming.*

Balland, P.A., R. Boschma & K. Frenken (2014). Proximity Dynamics: From Statics to Dynamics. Regional studies 49(6), pp, 907-920.

Balland, P.A. & D.L. Rigby (2017) The Geography of Complex Knowledge, Economic Geography, 93(1), pp. 1-23.

Bathelt, H., A. Malmberg & P. Maskell (2004). Local Buzz, Global Pipelines and the Process of Knowledge Creation. Progress in Human Geography 28(1), pp. 31-56.

Boschma, R. (2005). Proximity and innovation: a critical assessment. Regional Studies 39(1), 61–74.

Boschma, R. (2017). Relatedness as driver of regional diversification: a Research Agenda. Regional Studies 51(3), pp. 351-364.

Boschma, R., P.A. Balland & F.D. Kogler (2014). Relatedness and technological change in cities: the rise and fall of technological knowledge in U.S. metropolitan areas from 1981 to 2010. Industrial and Corporate Change 24(1), pp. 223-250.

Boschma, R. & K. Frenken (2009). Technological Relatedness and Regional Branching. In: Bathelt, Feldman and Kogler (eds.). Dynamic Geographies of Knowledge Creation and Innovation. Routledge: Taylor and Francis, pp. 1- 16.

Boschma, R. & K. Frenken (2010). The spatial evolution of innovation networks: a proximity perspective. In Boschma and Martin (eds),.The Handbook on Evolutionary Economic Geography. Cheltenham, UK and Northampton, MA: Edward Elgar, pp. 120–135.

Boschma, R. & C. Gianelle (2014). Regional branching and Smart Specialization Policy. S3 Policy Brief Series n° 06/2014.

Bottazzi, L & J. Peri (2003). Innovation and Spillovers in Regions: Evidence from European Patent Data. European Economic Review 47, pp. 687-710.

Breschi, S., F. Lissoni & F. Malerba. (2003). Knowledge-relatedness in firm technological diversification. Research Policy 32(1), pp. 69-87.

Broekel, T., P.A. Balland, M. Burger & F. van Oort (2014). Modelling Knowledge Networks in Economic Geography. A Discussion of four Methods. The annals of regional science 53(2), 423-452.

Broekel, T. (2014). The Co-evolution of Proximities - A Network Level Study. Regional Studies: The Journal of the Regional Studies Association 49(6), pp. 921-935.

Broekel, T. (2017). Measuring Technological Complexity - Current Approaches and a new Measure of Structural Complexity. https://arxiv.org/abs/1708.07357, pp. 1-37.
Cassi L., & A. Plunket (2015). Research Collaboration in Co-inventor Networks: Combining Closure, Bridging and Proximities. Regional Studies 49(6), pp. 936-954.
Castaldi, C., K. Frenken, & B. Los (2015). Related variety, unrelated variety and technological breakthroughs : an analysis of US state-level patenting. Regional Studies, 49(5), 767-781.
Cohen, W.M. & D.A. Levinthal (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. Administrative Science Quarterly 35(1), pp. 128-152.
Crespo, J., P.A., Balland, R. Boschma & D. Rigby (2017). Regional Diversification Opportunities and Smart Specialization Strategies. Luxembourg: Publications Office of the European Union pp. 1-26.
Dobusch, L. & E. Schüßler (2013). Theorizing path dependence: a review of positive feedback mechanisms in technology markets, regional clusters, and organizations. Industrial and Corporate Change 22(3), pp. 617–647.
Duranton, G. & Puga, D. (2001). Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products. The American Economic Review 91(5), pp. 1454-1477.
Van Eck, N.J. & L. Waltman (2009). How to normalize Cooccurrence Data? An Analysis of some well-known Similarity Measures. Journal of the Association for Information Science and Technology 60(8), pp. 1635-1651.
Engelsman, E.C. & A.J.F. van Raan (1994). A Patent-based Cartography. Research Policy 23(1), pp. 1-26.
Fleming, L. & O. Sorenson (2001). Technology as a Complex Adaptive System: Evidence from patent pata. Research Policy 30, pp. 1019–1039.
Glaeser, E.L., H.D. Kallal, J.A. Scheinkman & A. Shleifer (1992). Growth in Cities. The Journal of Political Economy 100(6), pp. 1126-1152.
Glaeser, E.L. (2011). Triumph of the City: How our Greatest Invention makes us richer, smarter, greener, healthier, and happier. New York, NY: Penguin Press.
Griliches, Z. (1990). Patent statistics as economic indicators: a survey, Journal of Economic Literature, Vol. 28(4), pp. 1661-1707.
Gross, T. & B. Blasius (2008), Adaptive coevolutionary networks: a review, J. R. Soc. Interface 5(20), pp. 259–271.
Hall, B.H., A.B. Jaffe and M. Trajtenberg (2001). The NBER Patent CItation Data File: Lessons, Insights and Methodological Tools. NBER Working Paper 8498, pp. 1-74.
Hidalgo, C.A., B. Klinger, A.L. Barabási & R. Hausmann (2007). The Product Space conditions the Development of Nations. Science 317(5837), pp. 482-487.
Hidalgo, C.A. & R. Hausmann (2009). The Building Blocks of Economic Complexity. Proceedings of the National Academy of Sciences of the United States of America 106(26), pp. 10570-10575.
Hipple, E. von (1994). "Sticky Information" and the Locus of Problem Solving: Implications for Innovation. Management Science 40(4), pp. 429-439.
Jacobs, J. (1969). The Economy of Cities. New York: Random House.
Jaffe, A.B. (1986). Technological opportunity and spillovers of R&D: evidence from firms' patents, profits and market value. National Bureau of Economic Research, Working Paper 1815. Cambridge: Massachusetts.
Jaffe, A.B., M. Trajtenberg & R. Henderson (1993). Geographic Localization of Knowledge Spillovers as evidenced by Patent Citations. The quarterly Journal of Economics 108(3), pp. 577-598.
Kauffman (1993). The Origins of Order: Self-Organization and Selection in Evolution. Oxford University Press: New York.
Kauffman, S. & W. Macready (1995). Technological Evolution and Adaptive Organizations. Ideas from Biology might find Applications in Economics. Complexity 1(2), pp. 26-43.
Koenker, R.W. (2004). Quantile Regression for Longitudinal Data. Journal of Multivariate Analysis 91(1), 74–89.
Koenker, R.W. & G.W. Bassett (1978). Regression quantiles. Econometrica 46(1), 33-50.
Koenker R. & Hallock K. F. (2001). Quantile Regression. Journal of Economic Perspectives 15(4), pp. 143–156.
Koenker, R.W. (2018). Package ‘Quantreg’. Cran R-project.org.
Kogler, F.D., D. Rigby & I. Tucker (2013). Mapping Knowledge Space and Technological Relatedness in US Cities. European Planning Studies, 21(9), pp. 1374-1391.

Kogler, F.D., J. Essletzbichler & D.L.Rigby (2017). The Evolution of Specialization in the EU15 Knowledge Space.
Journal of Economic Geography, Oxford University Press 17(2), p. 345-373.
Levin, S.A. (1998). Ecosystems and the Biosphere as Complex Adaptive Systems. Ecosystems 1, pp. 431-436.
Marshall, A. (1980). Principles of Economics. London: Macmillan and Co.
Martin, R. and Sunley, P. (2006). Path Dependence and Regional Economic Evolution. Journal of Economic Geography 6(4), pp. 395-437.
Fagiolo, G. & M. Mastrorillo (2014). Does Human Migration Affect International Trade? A Complex-Network Perspective. PLoS One 9(5), pp. 1-20.
Menzel, M.P. (2008). Dynamic Proximities - Changing Relations by Creating and Bridging Distances. Papers in Evolutionary Economic Geography #08.16, pp. 1-27.
Morgan, K. (2004). The exaggerated death of geography: learning, proximity and territorial innovation systems, Journal of Economic Geography 4, pp. 3–21.
Neffke, F., M. Henning & R. Boschma (2011). How do regions diversify over time? Industry relatedness and the development of new growth paths in regions. Economic Geography 87(3), pp. 237- 265.
Nelson, R.R. & S.G. Winter (1982). An Evolutionary Theory of Economic Change, Harvard University Press, Cambridge, MA.
Orlando, M.J. (2004). Measuring Spillovers from Industrial R&D: On the Importance of Geographic and Technological Proximity. The RAND Journal of Economics 35(4), pp. 777-786.
Pavitt, Keith (1988). Uses and Abuses of Patent Statistics. in: A. F. J. van Raan (eds). Handbook of Quantitative Studies of Science and Technology. Amsterdam: Elsevier Science Publishers.
Petralia, S., P.A. Balland & D.L. Rigby (2016). Histpat Dataset. Scientific Data 3, 160074. Harvard Dataverse: Cambridge.
Petralia, S., P.A. Balland & A. Morrison (2017). Climbing the Ladder of Technological Development. Research Policy 46(5), pp. 956 - 969.
Pintea M. & P. Thompson (2007). Technological Complexity and Economic Growth. Review of Economic Dynamics 10(2), pp. 276-293.
Ponds R., F. van Oort & K. Frenken (2007). The Geographical and Institutional Proximity of Research Collaboration. Papers in Regional Science 86(3), pp. 423-443.
Saviotti, P.P. & A. Pyka (2013). The Co-evolution of Innovation, Demand and Growth. Economics of Innovation and New Technology 22(5), pp. 461-482.
Scott, A.J. and M. Storper (2003). Regions, globalization, development. Regional Studies 37(6). pp. 579-593.
Schumpeter, J.A. (1942). Capitalism, Socialism, and Democracy. New York: Harper and Brothers.
Storper, M. (1995). The Resurgence of Regional Economies, Ten Years Later: The Region as a Nexus of Untraded Interdependencies. European Urban and Regional Studies 2(3), pp. 191-221.
sorenson, O., J.W. Rivkin and L. Fleming (2006), ‘Complexity, networks and knowledge flow’, Research Policy, 35
(7), pp. 994–1017.
Strumsky, D., J. Lobo & S. van der Leeuw (2012). Using Patent Technology Codes to Study Technological Change, Economics of Innovation and New Technology, 21(3), pp. 267-286.
Ter Wal A. L. J. (2013). The dynamics of inventor networks in German biotechnology: geographic proximity versus triadic closure. Journal of Economic Geography 14(3), pp. 589-620.
Torre A. and Rallet A. (2005). Proximity and localization, Regional Studies 39(1), 47–59.
United States Census Bureau (2018). https://www.census.gov/geo/reference/gtc/gtc_cbsa.html.
USPTO (2018). http://patents.reedtech.com/classdata.php
Wagner, A. & W. Rosen (2014). Spaces of the Possible: Universal Darwinism and the Wall between technological
and biological Innovation. Journal of the Royal Society Interface 11(97), pp. 2-11.
Weissing, F.J., P. Edelaar & G.S. van Doorn (2011). Adaptive Speciation Theory: a Conceptual Review. Behavioural Ecology and Sociobiology 65(3), pp. 461-480.

Appendix

Table 2: Alternative Specifications

	$\Delta Relatedness_{ij,t}$				$\Delta Co-location_{ij,t}$			
$AGE(t)$	.290***	.321***	.328***	.290***	-.005	-.073***	-.076***	-.072***
$\Delta Relatedness_{ij,t}$	-	-	-	-	.008***	.006***	.007***	.007***
$\Delta Co-location_{ij,t}$	.060***	.062***	.126***	.126***	-	-	-	-
$\Delta Relatedness_{ij,t-1}$	-.112***	-.114***	-.126***	-.126***	.004***	.004***	.003***	.003***
$\Delta Co-location_{ij,t-1}$	.023***	.026***	.065***	.065***	-.338***	-.442***	-.434***	-.434***
$Co-location_{ij,t-2}$	-	.014	3.297***	3.166***	-	-16.592***	-15.739***	-15.731***
$Relatedness_{ij,t-2}$	-	-.028***	-.028***	-.028***	-	.044***	.030***	.030***
$COMij * AGE(t)$	-	-	-	.082***	-	-	-	-.010
$COMij * \Delta Relatedness_{ij,t}$	-	-	-	-	-	-	-.002	-.002
$COMij * \Delta Co-location_{ij,t}$	-	-	-.088***	-.087***	-	-	-	-
$COMij * \Delta Relatedness_{ij,t-1}$	-	-	.024***	.024***	-	-	.001	.001
$COMij * \Delta Co-location_{ij,t-1}$	-	-	-.064***	-.062***	-	-	-.022***	-.023***
$COMij * Co-location_{ij,t-2}$	-	-	-5.688***	-5.299***	-	-	-2.616***	-2.679***
$COMij * Relatedness_{ij,t-2}$	-	-	-.007*	-.011***	-	-	.061***	.062***
Loglikelihood	5710	5810	5959	5965	15389	21384	21651	21652
Degrees of freedom	156	158	164	165	156	158	164	165