Technological Relatedness and Knowledge Space:
Entry and Exit of U.S. Cities from Patent Classes

DAVID L. RIGBY

Department of Geography, University of California Los Angeles, Los Angeles, CA 90095, USA
Email: rigby@geog.ucla.edu

U.S. patent and citation data are used to measure technological relatedness between major patent classes in the USPTO. The technological relatedness measures, constructed as the probability that a patent in class j will cite a patent in class i, form the links of a knowledge network. Changes in this knowledge network are examined from 1975 to 2005. Evolution of the patent knowledge base within U.S. metropolitan areas is tracked by combining the knowledge network with annual patent data for each city. Entries and exits of cities from patent classes are linked to local and non-local measures of technological relatedness.

knowledge space technological relatedness patents citations entry exit

JEL classifications: O18, O31, R12
INTRODUCTION

Regardless of whether the theory of growth is guided by political economy or more orthodox economic frameworks, the accumulation of knowledge and technological change are viewed as central to economic performance (SCHUMPETER, 1942; SOLOW, 1956; MARX, 1970; NELSON and WINTER, 1982; ROMER, 1986; LUCAS, 1988). For economic geographers, concern with the creation of competitive advantage at the regional level has long focused attention on the ability of place-based economic agents to acquire economically relevant knowledge and on their capacity to use that knowledge effectively (VON HIPPEL, 1988; COHEN and LEVINTHAL, 1990; LUNDVALL, 1992; STORPER, 1997; MASKELL and MALMBERG, 1999).

At least since the work of MARSHALL (1920), SCHUMPETER (1942) and PERROUX (1955), economic growth has been considered “lumpy”, or discontinuous, in both a spatial and a temporal sense. While geographical lumpiness might once have been seen as the outcome of localized natural resources, sunk capital investments and entrenched relationships between a region’s firms and industries, it is increasingly viewed as a product of the stickiness of tacit knowledge and the difficulty of developing the creative terroir between individuals, firms, and the panoply of supporting institutions from which knowledge is born (POLANYI, 1966; LUNDVALL, 1988; STORPER, 1995; MASKELL and MALMBERG, 1999). These ideas have spawned a massive literature on regional innovation systems (FREEMAN, 1985; COOKE et al., 1997), learning regions (MORGAN, 1997; LUNDVALL and JOHNSON, 1994) and localized knowledge economies more generally. MACKINNON et al. (2002) and ASHEIM and GERTLER (2005) offer reviews.
Though considerable effort has been directed towards uncovering what it takes to be a learning region or a knowledge economy, much less attention has been given to the character of knowledge produced within regions and how geographies of technology evolve over time. One of the primary reasons we know so little about the spatial composition of knowledge is that we lack precise measures of knowledge and technology (PAVITT, 1982). Consequently, researchers have long made use of proxies such as the “high technology” industry mix of a region’s economy (HALL and MARKUSEN, 1985), or the distribution of highly educated, skilled or creative “knowledge workers” (MARKUSEN et al. 2001; FLORIDA, 2002). Unfortunately, these proxies are noisy and they do not tell us much about knowledge created in different places.

This paper uses information contained within patents to trace the nature of knowledge created in U.S. cities over time. The arguments of the paper are laid out in four following sections. Section 2 provides additional motivation for this work and a brief review of related literature. Section 3 outlines a method of using patent citation data to measure the proximity of the different technology fields into which patents are placed. These measures form the links of a network that represents a U.S. knowledge space. Visualizations of that space are shown for the period 1975 to 2005. In Section 4, measures of proximity between technology fields are combined with metropolitan patent data to calculate the relatedness of knowledge production in U.S. cities. These measures are employed to model technological diversification and technological abandonment, key components of the evolution of urban knowledge cores. Section 5 concludes the paper.
LITERATURE REVIEW

Following SCHMOOKLER (1966) we often conceive of technology as knowledge of the industrial arts. In somewhat starker form, existent technology is equated with the known set of products and the processes used in their production (SAHAL, 1981). That parts of this set are more familiar to some economic agents than others leads readily to an understanding of technological specialization. This simple idea has endured as the basis for the division of labor (SMITH, 1776), for the industrial district (MARSHALL, 1920), through to competence based theories of the firm (WERNERFELT, 1984; BARNEY, 1991).

Economic geographers have long recognized geographical patterns of specialization in the distribution of industries (ELLISON and GLAESER, 1999), in techniques of production (RIGBY and ESSLETZBICHLER, 1997; 2006), and in organizational and institutional formations (SAXENIAN, 1994; STORPER, 1997). More recently, aggregation of these geographical differences has generated interest in the “knowledge-base” of regions (ALMEIDA, 1996; LAWSON and LORENZ, 1999; MASKELL and MALMBERG, 1999; CANTWELL and PISCITELLO, 2002; TODTLING and TRIPPL, 2005; MALECKI, 2007; QUATRARO, 2010). However, theoretical development of the concept of the regional knowledge-base has far outstripped empirical development. Beyond a few case studies, our empirical understanding of the geography of knowledge amounts to little more than inventories of activity within different economic categories across different locations. We know relatively little about the relationships between those categories, about the broad spatial structure of knowledge, or about the extent and the cohesion of the knowledge-base of specific places.

The first task of this paper, then, is to provide a new way of “mapping” knowledge space, and of identifying the position of regions within that space. A representation of U.S. knowledge
space is developed using patent data, based on the cognitive proximity of the 438 main technology classes identified in United States Patent and Trademark Office (USPTO) data. Regions are readily located within this broader knowledge space based on the frequency with which they generate patents of different kinds. This method is derived from measures of the range of firm knowledge portfolios (JAFFE, 1986) and the coherence of those portfolios (TEECE et al., 1994). Similar measures of technological or knowledge proximity are used by ENGELSMAN and VAN RAAN (1994) to “map” technological fields. BRESCHI et al. (2003), NESTA and SAVIOTTI (2005), LETEN et al. (2007) and COLOMBELLI et al. (2012) extend many of these arguments, offering different ways of calculating the structure of knowledge to explore technological diversification across firms and industries.

Attention has recently been focused on where new knowledge comes from. Building on the earlier claims of SCHUMPETER (1934), EVENSON and KISLEV (1976) offer a vision of applied research as a process of exploration within a space comprising existing ideas and all hybridized combinations of those ideas. WEITZMAN (1998) formally integrates this notion of invention as the recombination of past knowledge in a model of endogenous growth. Whereas for EVENSON and KISLEV (1976) diminishing returns to search within a finite knowledge space are assumed, WEITZMAN (1998) argues that new combinatorial possibilities are limited only by the processing capacity of agents. That capacity is conventionally regarded as bounded (SIMON 1959).

A recombinant vision of technology creation is insufficient as a model of localized, path dependent technological evolution. That model also rests on the following two arguments. First, the ability of firms to modify or discard existing technology is limited by the sunk costs of accumulating experience. From DAVID (1975), NELSON and WINTER (1982) and DOSI
(1982), we recognize that knowledge is accumulated largely through experimentation with existing techniques. Resources expended to develop and learn how to use technology take some time to recover, fixing firms within parts of knowledge space for some time. Second, the technological repertoire of most firms comprises only a small subset of the range of recombinant possibilities that define knowledge space, and there are real costs associated with search in that space (BINSWANGER, 1974). These costs are related to the topography of knowledge space that KAUFFMAN (1993) imagines as a fitness landscape where technologies are characterized by the number of components (ideas) and the extent of the interaction between them. Each of these technologies is associated with a level of fitness, an index of usefulness that agents can compare with their current technology. The level of autocorrelation, and thus the ease (cost) of search, within fitness landscapes is shown to depend on the extent of the interaction between the components that comprise particular technologies. OLSSON and FREY (2002) conjecture that the cost of recombinant search is a positive function of the technological distance between components of the knowledge set. This argument is exploited below. It is important to add that the structure of knowledge space itself evolves over time as new combinations of old and new ideas are developed. From related work (STUART and PODOLNY, 1996) we know that technological search is localized, a condition of sharply declining returns to investment in research and development efforts that are relatively dissimilar to existing technology, and by the costs of knowledge acquisition that rise steeply around the boundaries of existing knowledge bases (ATKINSON and STIGLITZ, 1969; WEBBER et al., 1992; ANTONELLI, 1995). The local character of search is thought by many to underpin the path dependent nature of much technological development (DAVID, 1985; ARTHUR, 1994).
The second main aim of this paper is to examine how the knowledge bases of regions shift over time. From a technological vantage point, the regional economy is a repository of specialized knowledge(s) that may be “located” in knowledge space. Where technologies from limited parts of knowledge space are associated with distinct geographical areas, localized communities of practice emerge (LAWSON and LORENZ, 1999; MASKELL and MALMBERG 1999; GERTLER 2003), reflecting place-specific sets of competences and capabilities (BOSCHMA and FRENKEN 2007). These capabilities reflect the legacy of past choices and they help shape the environment within which subsequent decisions are made (RIGBY and ESSLETZBICHLER 1997). Irreversibility and path dependence are critical features of the evolution of the space economy, with some regions “locked-in” to particular technological regimes that yield diminishing returns (GRABHER 1993), while others are better able to maintain their capacity to innovate (SAXENIAN, 1994), either through more open knowledge architectures or through connections to pools of knowledge generated elsewhere (BATHELT et al., 2004). Although these general arguments about geography and technology are reasonably well-accepted, at this time we know surprisingly little about the spatial structure of knowledge, about the important dimensions of that structure, and how these impact the evolution of technology in different places.

Recent empirical work has improved our understanding of the dynamic technological characteristics of firms and, in aggregate, industries and regions. In an innovative series of papers, FLEMING (2001) and FLEMING and SORENSON (2001) use patent data to model recombinant invention within a knowledge space that captures the component diversity and interdependence suggested by KAUFFMAN (1993). They establish the importance of interdependence and the size of the search space to the likelihood of successful search.
COLOMBELLI et al. (2013) extend the model of recombinant invention to define the structure of knowledge within the firm and report that firm survival is positively related to the coherence and variety of the firm’s knowledge base.

In more explicitly geographical, research VERTOVA (1999) and CANTWELL and VERTOVA (2004) illustrate the cumulative nature of technological development at the country-level, and the long-run relationship between country size and level of technological diversification. HAUSMAN and KLINERG (2007) and HIDALGO et al. (2007)) employ detailed trade data to measure the relatedness of products through patterns of co-exporting. They argue that specialization in the production of particular commodities provides countries with a set of capabilities that constrains diversification to related products. BOSCHMA et al. (2012) extend these ideas at the sub-national level exploring how different regions in Spain diversify into industrial sectors that are related to their existing product-based capabilities. COLOMBELLI et al. (2012) and BOSCHMA et al. (2013) define ‘technology spaces’ using patent data and illustrate how diversification at the local-level depends upon the existing structure of knowledge. Similar ideas are exploited by NEFFKE et al. (2011) to reveal path-dependence in the evolution of the industrial landscape in Sweden. QUATRARO (2010) shows that knowledge variety and coherence play a significant role in productivity growth across Italian regions, while KOGLER et al. (2013) illustrate linkages between technological relatedness and the pace of invention in U.S. cities. The more general concept of related variety is developed by FRENKEN et al. (2007) to explain differences in regional employment growth in the Netherlands.

The discussion of technological diversification and abandonment to this point assumes that cities are independent spatial units that are not influenced by technological practices elsewhere. Yet, a long history of geographical scholarship suggests that there are strong linkages
between urban areas (PRED, 1977). It seems reasonable therefore to extend the model of technological entry and exit to incorporate flows of knowledge between cities. Thus, economic agents within a city might gain knowledge of new technological combinations through development of those possibilities in other metropolitan areas. While JAFFE et al. (1993) illustrate the importance of geography to the flow of ideas, BRESCHI and LISSONI (2001) and SINGH (2005) suggest that social linkages maybe more important. FISCHER et al. (2006) explore the spatial as well as social limits on knowledge flows across EU regions, while FELDMAN et al. (2013) compare the influence of cognitive, geographical and social proximity on the diffusion of rDNA technology across U.S. cities. In the analysis below, the model of technological entry and exit that rests exclusively on the accumulation of local knowledge is extended to incorporate flows of technological information through networks of co-inventors spread across U.S. metropolitan areas.

THE U.S. KNOWLEDGE SPACE

Increasingly, we have turned to various measures of the inputs and outputs of invention and innovation to track knowledge production. On the input side, research and development spending has been shown to be closely correlated with counts of patents and innovations that typically form our indicators of knowledge outputs (FELDMAN, 1994). Many of these linkages are exploited by work on knowledge production functions and spatial derivatives thereof (GRILICHES, 1979; ACS et al., 2002). Patent data have become a staple for those interested in the geography and history of knowledge production (LAMOREAUX and SOKOLOFF, 1996; O’HUALLACHAIN, 1999; JAFFE and TRAJTENBERG, 2002; O’HUALLACHAIN and LEE, 2010), in inventors and inventor networks (BRESCHI and LISSONI, 2001; SINGH, 2005), in
knowledge flows or spillovers (JAFFE et al., 1993; SONN and STORPER, 2008), in geographical and cognitive proximity (FISCHER et al., 2006), and on the types of knowledge produced (HALL et al., 2001; STRUMSKY et al., 2012).

The popularity of patent data is related to their availability and to the wealth of information that they provide. At the same time, the disadvantages of patent statistics as overall indicators of economic and inventive activity are legion (PAVITT, 1985; GRILICHES, 1990), and we are becoming increasingly aware of the difficulties of extracting increasingly sophisticated information from patent and citation records (BRESCHI and LISSONI, 2005; THOMPSON and FOX-KEANE 2005; ALCACER and GITTELMAN 2006). Bearing these difficulties in mind, and recognizing that patents do not represent all knowledge produced, the primary aim of this section of the paper is to develop a measure of technological relatedness between patent classes, to use this measure to define a U.S. knowledge space, and to explore how this space has changed over time. The patent data are drawn from the United States Patent and Trademark Office (USPTO).

Upon review, individual patents are placed into one or more distinct technology classes that are supposed to reflect the technological characteristics of the underlying knowledge base that they embody. By the end of 2009, there were 438 such classes of utility patents in use by the USPTO. It is important to note that these technological classes do not remain constant over time. Through its bi-monthly “classification orders” the USPTO redefines classes, it adds new classes and, though rare, removes obsolete ones. Fortunately, the USPTO also reclassifies patents, providing the researcher with a consistent set of technology classes into which patents are placed for specific periods of time. Patents may be placed into a number of different technology classes, consistent with the range of knowledge that they introduce, though each granted patent is also
allocated a primary technology class on the basis of the extent of the novelty generated across different technology fields. The research below focuses upon these primary technology classes.

Attempts to generate measures of the relationships between patent technology classes can be traced back to JAFFE (1986), though SCHERER (1982) and TEECE et al. (1994) follow similar lines of argument in their measures of inter-industry R&D relatedness and corporate coherence, respectively. JAFFE (1986), VERSPAGEN (1997), BRESCHI et al. (2003) and others measure the technological proximity between different patent classes by examining the joint classification of individual patents across those classes. The methodology adopted here follows LETEN et al. (2007) using patent citations and the primary technology classes into which citing and cited patents are placed in order to measure the distance between them. In terms of mapping the structure of knowledge, this technique appears more direct than applying the method of HIDALGO et al. (2007) to find the conditional probabilities with which technology classes are found across a set of regions. The co-location of patents from different technology classes might be a reasonable proxy for technological relatedness, but might also be a proxy for unspecified economic relationships that display positive spatial autocorrelation.

To construct a U.S. knowledge space the location of individual patents has to be determined. For single inventor patents this is straightforward. For patents with multiple inventors, the country and sub-national region of the first listed inventor is taken as the location of the patent. The analysis here focuses upon U.S. inventions only, those where the first-listed inventor is located in the United States. Using the location of the first inventor, in cases of co-invention, assumes that the order of inventors reflects their relative weight in the invention process. Splitting patents across cities based upon the distribution of co-inventors made little difference to the results reported here. Only citations between U.S. patents are tracked.
The period of analysis runs from 1975 to 2005. 1975 is the first year for which patents are electronically linked to citations in the USPTO database. Patent data are available through 2011, though analysis here ends in 2005 largely because of the right censoring of patent applications. Note that the patent data are not averaged across a number of years: they are not smoothed in a temporal sense. There are significant swings in patent numbers and citations across individual years, but the years shown are certainly not outliers. Patents are aged by date of application rather than year of granting, for the usual reasons. Note that U.S. knowledge networks were built using both year of application and year of granting. There is little difference to the overall results. Self-citations are not removed in construction of the knowledge network for the analysis here is not attempting to identify spillovers. Rather, the focus is on the technological linkages across all patent classes. Investigation focuses on citations to patents that are no more than twelve years old. The twelve year cutoff is employed to ensure a long period of investigation while also capturing the majority of the citations that most patents generate (see HALL et al., 2001).

Table 1 shows the number of granted domestic patent applications to the USPTO for 1975, 1985, 1995 and 2005, along with the number of citations to existing U.S. patents recorded on those applications that are not more than twelve years old. For all these patents, citing and cited, their primary technological class is known. Clearly the number of U.S. patents has increased markedly over time. The apparent decline between 1995 and 2005 is, in large part, the result of right censoring, or truncation in the data: many patents applied for in 2005 have yet to be granted. Table 1 makes clear the rapid growth of total citations and the average number of citations per patent. Mean citations per patent climbed from an average of 2.8 in 1975 to 9.6 in
2005. The extent to which this represents citation inflation or the increasingly derivative nature of new knowledge claims is unclear.

TABLE 1 ABOUT HERE

The citation and technology class information for the patents described in Table 1 are used to derive a measure of the technological relatedness between all pairs of patent classes. In a general sense, two technology classes are considered related if patents in one of these classes cite patents in the other class. There are 438 primary utility patent technology classes currently used by the USPTO. All patents in the database are located in one of those primary classes.

A method for calculating technological relatedness between patent classes is outlined next. All (granted) patent applications for a given year, say 1975, are recorded along with all citations on those citing patents that extend back for twelve years. This generates a database of cited patents that extend back to 1963 (for the digital records). The primary technology classes of all citing and cited patents are recorded and arranged in the following matrix

\[
\begin{bmatrix}
    c_{11} & \cdots & c_{1438} \\
    \vdots & \ddots & \vdots \\
    c_{4381} & \cdots & c_{438438}
\end{bmatrix}
\]

where \( C_{ij}^t \) is a 438x438 matrix, the elements of which record the number of citations made by citing patents in technology class \( j \) to cited patents in class \( i \) in a given year \( t \). Dividing each element of \( C_{ij}^t \) by the number of patent applications (granted) in the element’s column class yields a matrix of the relative frequency that a patent in technology class \( j \) in a specified year will cite a patent in technology class \( i \)

\[
P_{ij}^t = \frac{c_{ij}^t}{N_j^t}
\]
where $N_j^t$ is the number of patents in technology class $j$ in a given year and $P_{ij}^t$ provides a measure of the technological relatedness or knowledge relatedness between patents in technology classes $i$ and $j$. The elements of $P_{ij}^t$ take the value 0 when patents in class $j$ do not cite patents in class $i$. In this case there is no technological relatedness between class $i$ and class $j$. To be more concrete, in 1985, technology class 331 (oscillators) did not cite technology class 236 (automatic temperature and humidity regulation). However, note that technology class 236 did cite technology class 331 in 1985. Thus, the matrix of relatedness between patent classes is asymmetric. The values of technological relatedness are not bounded to the right. Patents in the same technological class, located on the principal diagonal of $P_{ij}^t$, quite often exhibit relatedness values greater than 1. On average, technological relatedness should be greater for patents in the same technology class than for patents located in different classes. The values on the principal diagonal vary, perhaps, with the technological heterogeneity of patents found within individual classes. Technological relatedness values greater than 1 are rare off the principal diagonal of the matrix $P_{ij}^t$. A similar method of identifying technological relatedness was applied to European patent and citation data by LETEN et al. (2007). That analysis focused only on 30 technology classes and thus the technological precision of the estimates developed here is greater.

Individual technology classes reported in the USPTO have been aggregated into 30 intermediate classes and 6 broad technology classes by HALL et al. (2001). These aggregate groups afford a simple test of the efficacy of the technological relatedness measure just outlined. It makes sense to assume that primary patent classes that are members of the same aggregate technology groups will cite one another with greater frequency than primary patent classes found in different aggregate technology groups. If we do not see this, then the meaning of our relatively frequency matrix $P_{ij}^t$ is in doubt.
With the aid of UCINET (BORGATTI et al., 2002), the network of technological relatedness across the 438 primary patent classes is mapped. The network is generated with the Gower-scaling metric, itself derived to examine patterns of similarity across network nodes (GOWER, 1971). The nodes in the network correspond to each of the 438 distinct technological classes within the USPTO. The relative positions of the nodes are fixed by the frequencies of citation across each technology class pair \( P_{ij}^t \). The principal diagonal plays no role in the relative locations of the nodes. Note that because of the asymmetry in the relatedness matrix, the links between network nodes \( i \) and \( j \) are the average of \( P_{ij}^t \) and \( P_{ji}^t \). The knowledge relatedness networks for 1975-2005 are shown in Figure 1.

FIGURE 1 ABOUT HERE

The node colors in the figures represent the aggregate technology (6 class) grouping of HALL et al. (2001). There is clear evidence of the clustering of individual patent categories within most of these classes, indicating that the relatedness measure is capturing what may be considered as a common knowledge base within these more aggregate technology groupings. Network links are not reported for their density would render the network largely unreadable. However, note that because of the increase in the number of citations per patent throughout the study period the density of links increases markedly between 1975 and 2005. The citation data used to build the knowledge networks are not adjusted for potential citation inflation. The size of each node illustrates the number of patents in that technology class in the given year. Node sizes have been scaled to allow comparison over time in Figure 1.
The U.S. knowledge space in 1975 shows that patents appear to be reasonably evenly distributed across the six broad technology groups. The shared knowledge cores of those groupings can be identified in 1975 though the individual clusters become much more pronounced thereafter. The growth of drug and medical patents in the 1980s and 1990s is clear and the emergence of electronics and computers and communications technologies since the 1980s is also apparent. Figure 1 makes clear the rapid expansion in the share of patents in these three classes.

To measure the overall technological coherence of the patent network an average relatedness score is generated. Average relatedness measures the total “technological distance” between all pairs of patents divided by the number of such pairs. For a given number of patents, a higher average relatedness score indicates that patents are located in technology classes that are relatively close to one another in the U.S. knowledge network. These are patents found in classes that tend to cite each other with a relatively high frequency. A lower relatedness score would indicate that the patents are distributed over technology classes that are, on average, further apart from one another in the knowledge space. Average relatedness measures the technological specialization or coherence of produced knowledge. Higher levels of relatedness indicate greater technological specialization. The average relatedness value for a region $r$ in year $t$ is calculated as:

$$AR^{t,r} = \frac{\sum_i \sum_j P_{ij}^t \times D_{ij}^{t,r} + \sum_i P_{ii}^t \times 2D_{ii}^{t,r}}{N^{t,r} \times (N^{t,r} - 1)}$$

for $i \neq j$

where $P_{ij}^t$ measures the technological relatedness between patents in the 438 technology classes $i$ and $j$, $N^{t,r}$ is a count of the total number of patents in region $r$ in year $t$, and where $D_{ij}^{t,r}$ counts the number of pairs of patents that can be located in technology classes $i$ and $j$ in region $r$ in year $t$. 

To clarify the meaning of $D_{ij}^{t,r}$, imagine a region with three patents, one in technology class 1 and two in technology class 2. Then, the pair counts $D_{ij}^{t,r}$ represent elements in the (438x438) symmetric matrix

$$
D^{t,r} = \begin{bmatrix}
0 & 2 & \ldots & 0 \\
2 & 1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 0
\end{bmatrix}
$$

With three patents, there are $3 \times 2 = 6$ unique distance measures to calculate, the distance between the patent in class 1 and each of the patents in class 2, the distances from both patents in class 2 to the patent in class 1 and the distance between the two patents in class 2. Note that the latter distance is counted twice.

Tables 2 and 3 provide information on average technological relatedness between all patents in the U.S. knowledge network and between patents within each of the six aggregate technology classes. Table 2 reports that average relatedness increased by more than 94% between 1975 and 2005, even after adjusting for “citation inflation”. An increase in average relatedness indicates that more patents are being generated within technology classes that are closer to one another in technology space. This is consistent with the growth of technological specialization, an increase in the shared knowledge base that underpins invention. The rate of growth of specialization in U.S. patenting accelerated sharply after 1985, though it slowed somewhat between 1995 and 2005.

TABLE 2 ABOUT HERE

TABLE 3 ABOUT HERE
As expected, Table 3 shows that average relatedness values are much greater within each of the six aggregate technology classes than overall. This confirms expectations that technological information (a citation) is more likely to flow within a major technology grouping (such as the chemicals category) than it is to flow between such groupings. The drugs and medical group exhibits the highest average relatedness score of all major patent groups, indicating that knowledge, in the form of citations, circulates more frequently in this group than in others. On average, relatedness scores are relatively high in the electronics and in the computers and communication patent groups. They tend to be lower for chemical and mechanical patents. Since 1975, before adjusting for citation inflation, average relatedness values have increased fastest in the electronics, the drugs and medical, and the miscellaneous patent groups.

TECHNOLOGICAL DIVERSIFICATION AND ABANDONMENT IN U.S. METROPOLITAN AREAS

The knowledge cores of metropolitan areas in the U.S. can be identified by mapping the patents generated within individual cities onto the U.S. knowledge space of a given year. Building a separate knowledge space for each city based upon localized patent citations would yield very sparse networks for most because of the tendency for the majority of citations to cross metropolitan boundaries. What does the structure of a city’s knowledge base look like, and how do the knowledge bases of cities change over time? With 366 metropolitan areas, housing well over 90% of all U.S. patents, it is only possible to provide illustrations of the knowledge cores of a few. Three CBSAs are highlighted - Boise, Idaho, Dayton, Ohio and Rochester, New York. These cities do not form a representative sample, they serve merely to illustrate quite different
trajectories of knowledge production. Rather than map every patent within these three cities, only patent classes in which cities exhibit relative specialization, identified by location quotients greater than one, are displayed. Use of the location quotient serves to reduce the significance of individual patents that might be assumed to randomly populate the map of U.S. invention.

In 1975, Boise was home to inventors who developed 22 patents. 21 of these 22 patents were developed in different patent classes, with only one USPTO patent technology class accumulating two patents. Figure 2a shows the relative position of these 21 patent classes in the U.S. technology space for 1975. There is not much in this figure that suggests an existing, or even a nascent, technology or knowledge core. The average relatedness score for Boise in 1975, the average technological distance between all pairs of patents (not just those in classes with a location quotient greater than 1) was a relatively low 0.0123, ranking Boise 275 out of 366 metropolitan CBSAs in terms of the coherence of its knowledge base. However, by 2005, Boise had become one of the leading centers of semi-conductor patenting in the United States. In that year, Boise inventors generated 577 patents, 68% of these in just three technology classes: 257 (active solid state devices), 365 (static information storage and retrieval) and 438 (semi-conductor device manufacturing). Of 74 metropolitan CBSAs in 2005 generating more than 100 patents, Boise ranked second highest in terms of average relatedness (1.413), indicating that most of its patents were located in technology classes closely related to one another. While a detailed history of invention in Boise is well beyond the scope of this paper, the interested reader might consult MAYER’S (2011) account of the evolution of invention in Treasure Valley.
A very different pattern of urban invention is offered by Dayton, Ohio that traces a long history of aircraft, computing and automobile related patenting, through the work of the Wright brothers, the National Cash Register Company (NCR) and Dayton Engineering Laboratories Company (DELCO), later to become the foundation of the General Motors research arm. One of the most productive U.S. cities for patenting at the end of the nineteenth century, by the 1970s the pace of invention in Dayton had already begun a slow decline. Nonetheless, in 1975 with 289 patents, Dayton ranked 30 out of all U.S. cities in terms of the number of patents produced. These patents were distributed across mechanical, chemical and electronic technologies with one pronounced cluster in refrigerants (see Figure 2b). In 1975 the average relatedness of Dayton’s patents was 0.027, ranking 33 out of 60 metropolitan areas with more than 100 patents. As late as 2005, more than 100 patents were still produced in Dayton. However, these patents were widely scattered across the U.S. knowledge space. Indeed, the average relatedness score for Dayton’s patents in 2005 was 0.0799, placing the city dead last, in terms of technological coherence, out of all U.S. metropolitan areas with more than 100 patents. The distributed character of Dayton’s patents in 2005 is clear in Figure 2b.

FIGURE 2b ABOUT HERE

Figure 2c maps the locations of patents developed in Rochester, New York in 1975 and 2005. Most noticeable from this figure is Rochester’s long concentration in optics related invention, chiefly a function of the activities of the Kodak and Xerox Corporations. FELDMAN and LENDELS (2010) provide an overview of the emergence and development of the optics industry in the United States. In 1975, Rochester generated 654 patents, marking the city as the
thirteenth most inventive in the U.S.. Even more remarkable is the concentration of these patents in technology classes strongly connected to the optics industry: 382 (image analysis), 399 (electrophotography) and 430 (radiation image chemistry). In 1975, Rochester had a higher average relatedness score (0.2495) than any other U.S. metropolitan area that housed more than 100 patents. By 2005, the average technological distance between Rochester’s patents had not changed by much. In this year, the average relatedness of Rochester’s patents measured 0.2731, ranking the city 14th most specialized out of the most inventive U.S. metropolitan CBSAs. If the case of Boise represents the emergence of a knowledge core, Dayton provides an example of the dissolution of a knowledge core, and Rochester, NY exemplifies the maintenance of a regional knowledge base over time.

FIGURE 2c ABOUT HERE

How do existing configurations of technological capabilities within U.S. metropolitan areas shape trajectories of invention? This question is explored by using the structure of knowledge cores within cities to account for local histories of entry to and exit from individual patent classes. The goal is not to explain the precise patterns of technology production within individual cities, but rather to relate the technological structure of a place at time $t$ to the technological structure of that place, and its neighbors, in some future time $t+n$. Analysis focuses only on those patent classes (technologies) in which cities exhibit relative specialization. This is rendered, quantitatively, with a binary valued location quotient. Figure 3 illustrates the problem. In the left image, the shaded circles indicate those technologies in which a city is specialized in
year $t$, and the unshaded circles indicate those technologies in which a city is not specialized.

Which of the open circles is most likely to become shaded as we move from time $t$ to time $t+1$? That is, how does the knowledge core of a region diversify over time? In the right image of Figure 3, only those technologies in which the city has relative technological specialization are shown. The problem to solve in this image is which technology is most likely to be abandoned as we move from time $t$ to time $t+1$?

FIGURE 3 ABOUT HERE

A simple model assumes that technological diversification builds incrementally upon the existing knowledge base of the region. Thus, diversification to new technology classes, or gaining specialization in such classes, should be a function of their technological distance from the existing structure of knowledge within a region. In HAUSMANN and KLINGER (2007), diversification rests on the density of current practice within a product space and the value of that density around product classes that have yet to be exploited. I follow a similar logic and hypothesize that the probability of a city diversifying into a technology class is a positive function of the overall proximity (in knowledge space) of that class to all technology nodes in which the city is already specialized. In the left image of Figure 3, the city is most likely to diversify into knowledge class D because, out of the classes in which the city is not yet specialized, this class is closest (has maximum relatedness) to the set of techniques in which the city is already specialized. Along the same lines, it follows that cities will be most likely to abandon those technologies that are furthest from the core of their knowledge base. The right image of Figure 3 shows that of all technology nodes in which the city is specialized, node A is
the most remote, has the lowest total relatedness to all other specialized technology nodes in the city, and thus most likely to be abandoned.

Cities do not operate as independent economic units, rather they are imbricated in more or less dense webs of interaction that link economic actors across different locations. Information flows through these interactions perhaps signaling technological possibilities as yet untried in particular places. The greater the flow of information to a city, the more likely it is that the city’s knowledge base will be shaped by ideas developed elsewhere. To capture these flows, it is assumed that the probability of technological diversification in a city is influenced by the knowledge base (structure of knowledge) in other cities, together with an index of how closely those other cities are linked to the city in question. This is operationalized in the following way. For each year of data, a symmetric inter-city matrix (366x366) of co-inventor relations is multiplied by a (366x438) matrix of location quotients showing for each of the 366 U.S. metropolitan areas and 438 technology classes where each city has relative technological advantage (location quotient greater than 1). A given cell \((i, j)\) in the resulting (366x438) product matrix is the sum of the number of co-inventor pairs where one of the inventors is located in city \(i\) and the other inventor is located in a city other than \(i\) where the location quotient for technology class \(j\) is greater than 1. Thus, only flows of information from cities that have knowledge of a specific technology class are relevant for the entry of cities into that same technology class. The distance between one city and itself is set to zero, so I ignore intra-city co-inventor relationships. It is hypothesized that spillovers of knowledge between cities, driven by co-inventor relationships, should exert a positive influence on the probability that a city adopts a new technology.
To build the square matrix of co-inventor relations between all 366 metropolitan areas requires identification of individual inventors. The USPTO does not provide such data. Fortunately, LAI and colleagues at Harvard University have produced a list of individual inventors and their co-inventors that can be linked to the individual patent records in the USPTO (LAI et al., 2009). From these data I take all patent applications in a given year that list co-inventors and I record the metropolitan areas within which co-inventors are located. If a co-inventor was located outside the United States, or in one of the micropolitan CBSAs, they were dropped from the analysis. The metropolitan co-inventor matrix is initially populated with zero values in all cells. Patents with multiple inventors are then examined one-by-one. If two co-inventors on a patent are located in different metropolitan areas then the cells of the inter-city matrix of co-inventor relations corresponding to the cities where the co-inventors are located receive the value 1. The resulting matrix is symmetric. (Note that co-inventor counts along the principal diagonal measure the number of co-inventors located in the same city. This information is not exploited below.) If there are three co-inventors on a patent, each living in a different metropolitan area then six cells in the matrix receive a count of 1 (three pairs of cities in the matrix are linked and the symmetry ensures a count of six). This process is repeated for all co-invented patents with the inter-city matrix counts building a representation of the interaction between co-inventors located in different U.S. metropolitan areas in a given year.

The processes of metropolitan technological diversification and abandonment are examined using a panel version of a fixed effects logit model. The observational units are the 438 technology classes within each city over time. The model is run for every second year from 1975 to 2005 inclusive. The values of the dependent variable are 0 or 1, so the regression model is predicting the probability that $Y = 1$, that a city exhibits relative technological specialization in
a particular technology class in a given year. The binary nature of the dependent variable
suggests use of a probit or logit model extended to panel form to take advantage of the time
dimension in the data. This is not straightforward, for a probit model cannot be run with a fixed
effects panel specification that is suggested by a simple HAUSMAN test as preferable to a
random effects model. So I opt for a fixed effects panel version of the logit model.

The model to be estimated is

$$Y_{it}^c = \alpha + \beta_1 \text{Proximity}_{it-1}^c + \beta_2 \text{MSAnet}_{it-1}^c + z_t + city_c + tech_i + \epsilon_{it}^c$$

where the binary dependent variable assumes the value 0 or 1, and represents the probability of
city $c$ in year $t$ exhibiting relative technological specialization in technology class $i$. On the right
hand-side of the model, the first independent variable ($\text{Proximity}$) is the time-lagged value of the
total distance (in units of technological relatedness) between each technology class $i$ and all other
technology classes where the city exhibits relative technological specialization. The second
independent variable ($\text{MSAnet}$) is a time-lagged measure of co-inventor knowledge flows to city
$c$ from all MSA’s that have relative technological specialization in technology class $i$. The $z_t$
represent year fixed effects, $city_c$ and $tech_i$ are MSA and technology (patent) class fixed effects,
respectively.

Results are displayed in Table 4. Note that the models for entry restrict the observations
to those city-technology pairs in which the lagged value of the dependent variable is zero. The
models for exit restrict observations to those city-technology pairs in which the lagged value of
the dependent variable is 1. The conditional logits drop all city-technology groups in which the
value of the dependent variable is fixed over time. For purposes of comparison, results from a
linear probability model are also provided. In this model, standard errors are clustered at the city
level.
Table 4 provides clear support for the hypothesis that the pattern of technological diversification is guided by the existing knowledge structure within the city. The linear probability model suggests that a one unit increase in the proximity of a city’s knowledge base to a technology in which the city has no existing specialization increases the probability of developing that technological specialization by 0.7%. Table 4 also indicates that a city well-connected by co-inventor relationships to other cities that are specialized within a particular technology class is more likely to develop specialization in that same technology class than a city that is less well-connected. The size of this network effect is considerably smaller than the size of the local cognitive proximity effect. The partial regression coefficients for the conditional logit model yield consistent results, with both proximity and network effects being significant and positive.

Turning to the influence of the independent variables on technological abandonment or exit within the city, the results tell much the same story, though the impact of the city-network variable is insignificant in the linear probability model. For the case of exit, as the proximity of a technology class in which the city is specialized increases in relation to all other classes in which the city has specialization, so the probability of exit decreases. The linear probability model suggests that a one-unit increase in such proximity lowers the probability of exit by 2.72%. Thus, those technology classes in which a city is specialized, but which are located further away from the knowledge core of a city, have a greater chance of being abandoned. The logit model results are again consistent, indicating that cities are more likely to exit activity in technology classes that are remote from their knowledge cores and in which the cities that they are connected to through co-inventor relationships also are less active.
The results in Table 4 might be considered as somewhat preliminary in the sense that a broad set of time-varying covariates that might influence patterns of technological specialization within cities have not been added to the model. Given the data available at this time, two questions readily spring to mind. First, are the results in Table 4 robust to the length of the time periods across which observations are generated? Second, does use of a location quotient to indicate technological specialization exert any bias? These questions are answered in Table 5 where the models of entry and exit are estimated for 5-year periods 1975-79, 1980-84, …, 2000-04, and where specialization in a technology class is measured simply by its presence or absence within a city.

Table 5 offers qualified support for the results already presented. The development of city specialization into a particular technology class, from one 5-year period to the next, is a significant and positive function of the proximity of that class to the city’s existing knowledge core. This result holds whether specialization is measured using the binary-valued location quotient or simply by the presence of patents in the technology class. Cities are more likely to develop technological specialization in classes for which the cities they are well-connected to exhibit specialization, though this influence is only significant in the model where specialization is captured through the presence of patents in particular classes. Losing specialization (exit) in a city-technology pairing is a negative function of proximity to the city’s knowledge core and to the patterns of specialization in the city’s networked partners. However, note that the influence of proximity in the exit model without the location quotient measure of specialization is not
significant. In general, then, the loss of specialization in a technology class is more likely when that class is further from the knowledge core of the city in question and further from the knowledge cores of those metropolitan areas the city is connected to through co-inventor linkages. Note that adding one to all variables and log transforming them does not alter the basic findings reported above.

CONCLUSION
Citation data were used to measure the proximity of different technology classes into which patents are placed. The resulting measures of knowledge relatedness formed the edges of a patent network that maps the U.S. knowledge space. The evolution of that space was traced between 1975 and 2005. Over that time period, average relatedness between U.S. patents, after adjusting for citation inflation, has just about doubled: patents are increasingly concentrating in fewer technology classes and the distance between those classes is shrinking. Since 1975, the share of patents in chemical and mechanical classes has been decreasing while the share in drugs and medical, electronics, and computing classes has been increasing. Average relatedness scores vary markedly between these broad patent groupings.

While some cities transition extremely rapidly from one knowledge core to another, for most, the process of technological transition is relatively slow. Cities build competence around a range of related technologies over time and this competence shapes the knowledge trajectories that most cities tend to follow. Technological diversification in cities, an expansion of the knowledge core, depends upon current practice and the proximity of new technological possibilities to the set of existing specializations. Diversification is also influenced by information about knowledge production from other locations. Knowledge specialization
exhibits considerable inertia and the same forces that guide diversification play an even stronger role in maintaining competence. Technological abandonment is most likely to occur at the edges of the knowledge core occupied by a city.

Much more work remains to be done to define more precisely the knowledge cores of cities and how they evolve over time. Which technology classes are part of a knowledge core and which remain outside are important questions, along with identification of related knowledge sets that link different cores. What possible combinations of technologies are most productive, and how cities can efficiently transition from relatively barren parts of technology space to more fertile areas are key questions for future research. The importance of proximity in shaping the evolution of knowledge cores is clear, but what kinds of proximity are most important, how do structures of proximity vary, and how best should they be measured remain unknown.
Acknowledgements – The author thanks two anonymous referees who provided detailed comments that significantly improved the quality of this paper.

NOTES

1. While Kaufmann’s models feature links between technologies, these links do not reflect similarities in technologies but rather design interdependencies that force changes in linked parts to be coordinated.

2. The average relatedness values in Table 2 are calculated for all patents in the given year. The average relatedness values in Table 3 are calculated only for patents within the same aggregate technology grouping, after Hall et al. (2001).

3. Limiting the observations to those cases where the lagged value of the dependent variable is 1, the probability model is really examining retention, looking at the influence of the independent variables on the probability that the dependent variable remains equal to 1. Reversing the 0/1 binary values of the dependent variable focuses attention on exit. In this case Y=1 signifies exit of a city from a technology class.
REFERENCES


Table 1. Application year patents and grant year patents, US only

<table>
<thead>
<tr>
<th>YEAR</th>
<th>PATENTS APPLCTN. YEAR</th>
<th>CITATIONS (12-YEARS)</th>
<th>PATENTS GRANT YEAR</th>
<th>CITATIONS (12-YEARS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975</td>
<td>41,385</td>
<td>118,742 (2.817)²</td>
<td>45,692</td>
<td>116,237</td>
</tr>
<tr>
<td>1985</td>
<td>36,996</td>
<td>135,866 (3.604)</td>
<td>37,772</td>
<td>122,364</td>
</tr>
<tr>
<td>1995</td>
<td>80,460</td>
<td>549,062 (6.708)</td>
<td>54,670</td>
<td>276,495</td>
</tr>
<tr>
<td>2005¹</td>
<td>52,975</td>
<td>517,004 (9.649)</td>
<td>46,183</td>
<td>492,597</td>
</tr>
</tbody>
</table>

Notes: 1. 2005 application year patents are right censored. The patents identified here were granted by the end of 2009. Thus, after 2010 there were more than 52,975 granted U.S. patents with an application date of 2005. 2. Ratio of citations to patents in parentheses.
Table 2. Total and average knowledge relatedness, US total (all patents)

<table>
<thead>
<tr>
<th>YEAR</th>
<th>PATENTS(^1)</th>
<th>TOTAL</th>
<th>AVERAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975</td>
<td>41,385</td>
<td>23,811,636</td>
<td>0.01390</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(23,811,636)(^2)</td>
<td>(0.01390)(^2)</td>
</tr>
<tr>
<td>1985</td>
<td>36,996</td>
<td>28,041,166</td>
<td>0.02049</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(21,924,289)</td>
<td>(0.01602)</td>
</tr>
<tr>
<td>1995</td>
<td>80,460</td>
<td>366,530,344</td>
<td>0.05662</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(153,939,666)</td>
<td>(0.02378)</td>
</tr>
<tr>
<td>2005</td>
<td>52,975</td>
<td>259,254,666</td>
<td>0.09238</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(75,694,793)</td>
<td>(0.02697)</td>
</tr>
</tbody>
</table>

Notes: 1. Number of (granted) patent applications in year indicated, with known CBSA. 2. Knowledge relatedness adjusted for citation inflation in parentheses.
Table 3. Average knowledge relatedness by major patent class, US total

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CHEMICALS</td>
<td>0.05518</td>
<td>(0.05518)</td>
<td>0.07951</td>
<td>(0.06216)</td>
</tr>
<tr>
<td></td>
<td>0.11638</td>
<td>(0.04888)</td>
<td>0.13150</td>
<td>(0.03839)</td>
</tr>
<tr>
<td>COMPUTERS &amp; COMMNCTN.</td>
<td>0.14008</td>
<td>(0.14008)</td>
<td>0.14444</td>
<td>(0.11293)</td>
</tr>
<tr>
<td></td>
<td>0.24182</td>
<td>(0.10156)</td>
<td>0.26731</td>
<td>(0.07804)</td>
</tr>
<tr>
<td>DRUGS &amp; MEDICAL</td>
<td>0.18804</td>
<td>(0.18804)</td>
<td>0.33333</td>
<td>(0.26061)</td>
</tr>
<tr>
<td></td>
<td>0.75794</td>
<td>(0.31833)</td>
<td>0.92738</td>
<td>(0.27074)</td>
</tr>
<tr>
<td>ELECTRONIC</td>
<td>0.07994</td>
<td>(0.07994)</td>
<td>0.10972</td>
<td>(0.08578)</td>
</tr>
<tr>
<td></td>
<td>0.19388</td>
<td>(0.08143)</td>
<td>0.47144</td>
<td>(0.13764)</td>
</tr>
<tr>
<td>MECHANICAL</td>
<td>0.03508</td>
<td>(0.03508)</td>
<td>0.03966</td>
<td>(0.03101)</td>
</tr>
<tr>
<td></td>
<td>0.08184</td>
<td>(0.03437)</td>
<td>0.17182</td>
<td>(0.05016)</td>
</tr>
<tr>
<td>MISCELLANEOUS</td>
<td>0.04221</td>
<td>(0.04221)</td>
<td>0.05753</td>
<td>(0.04498)</td>
</tr>
<tr>
<td></td>
<td>0.10743</td>
<td>(0.04512)</td>
<td>0.25746</td>
<td>(0.07516)</td>
</tr>
</tbody>
</table>

Notes: 1. Knowledge relatedness adjusted for citation inflation in parentheses.
Table 4. Regression analysis of the probability of technological diversification and technological abandonment

<table>
<thead>
<tr>
<th></th>
<th>ENTRY</th>
<th>EXIT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>XTLPM</td>
<td>XTLOGIT</td>
</tr>
<tr>
<td>Proximity</td>
<td>0.0069*** (0.0004)</td>
<td>0.1111*** (0.0047)</td>
</tr>
<tr>
<td>MSAnet</td>
<td>0.00001*** (2.10e-06)</td>
<td>0.0001*** (0.0000)</td>
</tr>
<tr>
<td></td>
<td>n=2,128,216 n=531,681</td>
<td>n=77,826 N=54,656</td>
</tr>
<tr>
<td></td>
<td>LR=6530.97</td>
<td>LR=1342.87</td>
</tr>
<tr>
<td></td>
<td>χ² = 0.000</td>
<td>χ² = 0.000</td>
</tr>
</tbody>
</table>

Notes: XTLPM model fit with ordinary least squares. This is a panel version of the linear probability model with a binary dependent variable. Errors are clustered by cbsa. XTLOGIT is the conditional logit model estimated using maximum likelihood techniques. City, technology class and year fixed effects included in all models.
Table 5. Robustness checks for technological entry and exit

<table>
<thead>
<tr>
<th></th>
<th>ENTRY</th>
<th>EXIT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>XTLOGIT 5YR LQ</td>
<td>XTLOGIT 5YR NO LQ</td>
</tr>
<tr>
<td>Proximity</td>
<td>0.1816***</td>
<td>0.0792***</td>
</tr>
<tr>
<td></td>
<td>(0.0097)</td>
<td>(0.0077)</td>
</tr>
<tr>
<td>MSAnet</td>
<td>7.98e-06</td>
<td>0.0005***</td>
</tr>
<tr>
<td></td>
<td>(0.0000)</td>
<td>(0.0001)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>EXIT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>XTLOGIT 5YR LQ</td>
</tr>
<tr>
<td>Proximity</td>
<td>-0.2934*</td>
</tr>
<tr>
<td></td>
<td>(0.1622)</td>
</tr>
<tr>
<td>MSAnet</td>
<td>-0.00002**</td>
</tr>
<tr>
<td></td>
<td>(1.18e-06)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>EXIT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>XTLOGIT 5YR NO LQ</td>
</tr>
<tr>
<td>Proximity</td>
<td>-0.2467</td>
</tr>
<tr>
<td></td>
<td>(0.1998)</td>
</tr>
<tr>
<td>MSAnet</td>
<td>-0.0001***</td>
</tr>
<tr>
<td></td>
<td>(0.0000)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>ENTRY</th>
<th>EXIT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=123,339</td>
<td>n=94,920</td>
</tr>
<tr>
<td></td>
<td>LR=9,389.9</td>
<td>LR=21,914.4</td>
</tr>
<tr>
<td></td>
<td>$\chi^2 = 0.000$</td>
<td>$\chi^2 = 0.000$</td>
</tr>
<tr>
<td></td>
<td>n=27,734</td>
<td>n=39,759</td>
</tr>
<tr>
<td></td>
<td>LR=2,660.0</td>
<td>LR=11,520.0</td>
</tr>
<tr>
<td></td>
<td>$\chi^2 = 0.000$</td>
<td>$\chi^2 = 0.000$</td>
</tr>
</tbody>
</table>

Notes: The number of observations varies within the entry and exit models because patterns of specialization are different when calculated using presence/absence of a technology class in a city and when using the location quotient to identify specialization. LQ indicates specialization is represented by the binary-valued location quotient. NO LQ indicates specialization is represented by the presence/absence of the technology class.
Fig. 1. U.S. knowledge space

Notes: Black = Chemicals (1), Green = Computers & Communications (2), Yellow = Drugs & Medical (3), Red = Electronics (4), Blue = Mechanical (5), Grey = Miscellaneous (6). The largest node across the four years is USPTO class 514 in 1995, representing some 3219 patents.
Fig. 2a. The knowledge core of Boise

Notes: Figures 2a-2c are scaled independently to facilitate visualization. The sizes of nodes between cities or over time is not strictly comparable.
Fig. 2b. The knowledge core of Dayton

1975

2005
Fig. 2c. The knowledge core of Rochester

1975

2005
Fig. 3. Specialization of technology nodes in a city

Notes: The shaded circles indicate the technologies in which a city is specialized in relative terms.