Regional Studies

ISSN: 0034-3404 (Print) 1360-0591 (Online) Journal homepage: www.tandfonline.com/journals/cres20

rKnowledge: The Spatial Diffusion and Adoption of rDNA Methods

Maryann P. Feldman, Dieter F. Kogler & David L. Rigby

Regional Studies, 2015
Vol. 49, No. 5, 798–817, http://dx.doi.org/10.1080/00343404.2014.980799

Routledge
Taylor & Francis Group

Regional Studies Association

rKnowledge: The Spatial Diffusion and Adoption of rDNA Methods

MARYANN P. FELDMAN*, DIETER F. KOGLER† and DAVID L. RIGBY‡
**Department of Public Policy, University of North Carolina, Chapel Hill, NC, USA. Email: maryann.feldman@unc.edu*
†School of Geography, Planning & Environmental Policy, University College Dublin, Dublin, Ireland. Email: dieter.kogler@ucd.ie
‡Department of Geography, University of California, Los Angeles, CA, USA. Email: rigby@geog.ucla.edu

(Received February 2013; in revised form September 2014)

FELDMAN M. P., KOGLER D. F. and RIGBY D. L. rKnowledge: the spatial diffusion and adoption of rDNA methods, *Regional Studies*. The 1980 patent granted to Stanley Cohen and Herbert Boyer for their development of rDNA technology played a critical role in the establishment of the modern biotechnology industry. From the birth of this general-purpose technology in the San Francisco Bay area, rDNA-related knowledge diffused across sectors and regions of the US economy. Patent data are used here to track the geography and timing of rDNA technology adoption in US metropolitan areas. Using event history and fixed effects conditional logit models, it is shown how the diffusion of rDNA techniques was influenced by cognitive, geographical and social proximity.

Evolutionary economic geography Technology evolution Knowledge diffusion/adoption Patent analysis General-purpose technology rDNA methods

FELDMAN M. P., KOGLER D. F. and RIGBY D. L. r知识：rDNA方法的空间散佈与採用，*区域研究*。1980年授予史丹尼．科恩（Stanley Cohen）与赫博．玻伊尔（Herbert Boyer）所研发的rDNA技术之专利，在现代生物科技产业的发展中，扮演了关键的角色。从此一通用技术在加州湾区诞生开始， rDNA 的相关知识便扩散至美国经济的各部门及区域。我们运用专利数据，追溯美国大都会地区採用rDNA技术的地理与时间。本文运用事件历史与固定效应条件逻辑特模型，显示 rDNA 技术的散佈如何受到认知、地理与社会邻近性所影响。

演化经济地理 技术演化 知识散佈 / 採用 专利分析 通用技术 rDNA 方法

FELDMAN M. P., KOGLER D. F. et RIGBY D. L. La connaissance-r: la diffusion et l'adoption spatiales des méthodes ADNr, *Regional Studies*. Le brevet de 1980 délivré à Stanley Cohen et à Herbert Boyer en reconnaissance du développement de la technologie ADNr a joué un rôle essentiel dans l'établissement de l'industrie biotechnologique moderne. Depuis la naissance de cette technologie universelle dans la baie de San Francisco, la connaissance liée à l'ADNr s'est diffusée à travers des secteurs et des régions de l'économie des É-U. On emploie ici les données sur les brevets pour découvrir la géographie et le calendrier de l'adoption de la technologie ADNr dans les zones métropolitaines des É-U. Se servant de l'historique des événements et des modèles logit conditionnels à effets fixes, on montre comment la diffusion des techniques ADNr a été influée par la proximité cognitive, géographique et sociale.

Géographie économique évolutionniste Évolution technologique Diffusion/adoption de la connaissance Analyse des brevets Technologie universelle Méthodes ADNr

FELDMAN M. P., KOGLER D. F. und RIGBY D. L. rWissen: räumliche Diffusion und Einführung von rDNA-Methoden, *Regional Studies*. Für die Gründung der modernen Biotechnologiebranche spielte das Patent, das Stanley Cohen und Herbert Boyer 1980 für ihre Entwicklung der rDNA-Technik erhielten, eine entscheidende Rolle. Nach der Geburt dieser Basistechnologie in der San Francisco Bay Area diffundierte das Wissen über rDNA in verschiedene Sektoren und Regionen der gesamten US-Wirtschaft. In diesem Beitrag werden die Geografie und Zeitpunkte der Einführung von rDNA-Technik in den Ballungsgebieten der USA anhand von Patentdaten nachverfolgt. Mithilfe von Ereignisverlaufs- und konditionalen Festeffekt-Logitmodellen wird gezeigt, wie die Diffusion von rDNA-Techniken von kognitiver, geografischer und sozialer Nähe beeinflusst wurde.

Evolutionäre Wirtschaftsgeografie Technische Entwicklung Diffusion/Einführung von Wissen Patentanalyse Basistechnologie rDNA-Methoden

FELDMAN M. P., KOGLER D. F. y RIGBY D. L. Conocimiento-r: la difusión espacial y la adopción de los métodos ADNr, *Regional Studies*. La patente concedida en 1980 a Stanley Cohen y Herbert Boyer por su desarrollo de la tecnología de ADN recombinante desempeñó un papel fundamental en el establecimiento de la industria de la biotecnología moderna. Desde el nacimiento de esta tecnología básica en la región de la Bahía de San Francisco, el conocimiento relacionado con la ADNr se difundió en todos los sectores y regiones de la economía estadounidense. En este estudio utilizamos los datos de patentes para rastrear la geografía y el momento de la adopción de la tecnología ADNr en las zonas metropolitanas de los Estados Unidos. Con ayuda de modelos logit condicionales sobre acontecimientos de la historia y efectos fijos, mostramos cómo influyó la proximidad cognitiva, geográfica y social en la difusión de las técnicas de ADNr.

Geografía económica evolutiva Evolución tecnológica Difusión/adopción del conocimiento Análisis de patentes Tecnología básica Métodos ADNr

JEL classifications: L65, O00, O31, O33

INTRODUCTION

In December 1980, the United States Patent and Trademark Office (USPTO) issued a patent entitled *Process for Producing Biologically Functional Chimeras* (number 4237224). The patent covered the recombinant DNA (rDNA) technique developed by Dr Stanley Cohen of Stanford University and Dr Herbert Boyer of the University of California – San Francisco. In later evaluation of the Cohen–Boyer patent, the USPTO introduced a new category (435/69.1) to its classification system, a relatively rare occurrence signalling the birth of a new type of technology. While most technological innovation is incremental, certain discoveries provide fundamental breakthroughs that revolutionize industrial activity and provide a platform for increased productivity throughout the economy. The rDNA patent represented this kind of transformative general purpose technology laying the foundations for the growth of the modern biotechnology industry (FELDMAN and YOON, 2012).

From its conception in the Bay Area of California, knowledge of the Cohen–Boyer discovery spread rapidly to select cities across the country, stimulating the development of nascent clusters of biotechnology production. The diffusion of rDNA technology might have followed many different trajectories, but only one historical geography played itself out. This paper seeks to understand the forces shaping this geography in the characteristics of different metropolitan areas, in the knowledge structure of different cities, in the social ties that bind cities to one another through inventor collaborations, and in terms of the friction of distance.

Theory argues that innovative activity tends to cluster in regions where resources relevant to the performance and survival of knowledge-based firms are most abundant (FELDMAN, 1994), including the presence of skilled workers and labour market flexibility (SAXENIAN, 1994; GLAESER, 2000), proximity to markets and input suppliers (STORPER and CHRISTOPHERSON, 1987; BAUM and HAVEMAN, 1997), the presence of universities and research organizations (ZUCKER *et al.*, 1998), and cultural and institutional supports for entrepreneurial activity (SAXENIAN, 1994; SORENSON and AUDIA, 2000). Yet, the ways in which radical technological breakthroughs diffuse through space and the ways in which new scientific discoveries are adopted and incorporated with existing expertise have not been considered in a systematic way. We appeal for some guidance in this regard to the classical literature on the spatial diffusion of innovations (HAGERSTRAND, 1953; GRILICHES, 1957; BROWN, 1981) updated with more recent claims on the different dimensions of proximity (NOOTEBOOM, 2000; BOSCHMA, 2005), on absorptive capacity (COHEN and LEVINTHAL, 1990), and the importance of social relationships to knowledge flow (BRESCHI *et al.*, 2003; SORENSON, 2003; COWAN and JONARD, 2004).

Armed with such, this paper aims to shed some empirical light on the emergence of innovative, technology-driven industrial clusters. This is done by examining a specific technology and its diffusion over time and across space. That space takes different forms in this work: it is always geographical, but it is also constituted through sets of social relationships and the shifting character of regional knowledge bases. The focus is on Cohen–Boyer's rDNA technology because the boundaries of this new knowledge subset are relatively well defined by USPTO class 435/69.1 and because the geographical movement and adoption of rDNA technologies are captured in approximately 9000 patents distributed across US metropolitan areas. Working from an evolutionary economic geography framework, the influence of cognitive proximity, geographical proximity and social proximity on the diffusion and adoption of rDNA knowledge across US cities is explored. The analysis employs event history models to examine the timing of a city's first rDNA patent and the more general class of fixed effect panel models to explore the determinants of repeated rDNA technology adoption across the US urban system. Results show that the diffusion and adoption of the Cohen–Boyer technology was directed mainly by social contacts developed among co-inventors and by the absorptive capacity of host cities, measured through the structure of their knowledge cores and the cognitive proximity of those cores to knowledge in technology class 435/69.1. The initial diffusion of rDNA technology followed a hierarchical pattern jumping between relatively large and distant urban areas. A subsequent phase of diffusion sees rDNA technology radiating out from these larger cities exhibiting a classic distance–decay-type pattern. The results highlighted here appear robust to variations in model structure and to concerns with endogeneity.

The paper is organized in five sections. The second section offers a brief review of theory that links key ideas in evolutionary economic change to the diffusion and adoption of new technological knowledge. The third section provides more historical detail about the Cohen–Boyer patent along with a note on the structure of the USPTO classification system. The fourth section turns attention to the geographical and sectoral diffusion of rDNA technology after 1980 to a simple model specification, and to descriptive information about key independent variables in that model. The fifth section presents results from a Cox hazard model of rDNA technology diffusion extended to incorporate time-varying independent variables, along with output from a more general fixed effects logit model. The latter model is the platform for a series of robustness checks of the basic hypotheses. The sixth section offers some brief discussion and conclusions.

TECHNOLOGY EVOLUTION AND ADOPTION IN AN EVOLUTIONARY FRAMEWORK

The economic landscape is in constant flux, pushed and pulled by processes of competition that both trap capital, labour and routines within some sectors and regions while encouraging experimentation and discovery in others. Evolutionary economic geography offers a broad, though still contested, framework that has at its core the production and destruction of novelty in space and the links between novelty and regional economic fortunes (Boschma and Martin, 2010). The creation of technological knowledge, its movement and recombination within different regional ensembles of economic agents and institutions plays only one, but nonetheless critical, role in the evolution of the space-economy (Rigby and Essletzbichler, 1997).

It has become convention to characterize the history of technological change as comprising long periods of more or less constant incremental improvement punctuated by a periodic bursts of basic discovery, the creative gales of innovation that usher in new knowledge systems and that shift parts of the economy to new planes of development (Schumpeter, 1942; Nelson and Winter, 1982). The temporal lumpiness of basic innovation (Mensch, 1975) is mirrored by its uneven geography, with islands of innovation emerging from the economic landscape, sometimes remote and sometimes connected via heterogeneous social and economic networks, that bloom and wither as economic agents compete within, and simultaneously shape the evolution of, the capitalist space economy (Essletzbichler and Rigby, 2007).

Over the course of history there have been many attempts to identify and define technologies that are radical and to separate them from innovations that are more incremental in nature (Sahal, 1981; Dosi, 1982; Nelson and Winter, 1982; Abernathy and Clark, 1985; Clark, 1985). Recent research on the importance of inventions focuses on valuations of patents through forward citations (Trajtenberg, 1990a; Hall *et al.*, 2005, 2007; Gambardella *et al.*, 2008), though some are highly critical of this approach (Bessen, 2008; Abrams *et al.*, 2013). Interest in breakthrough inventions focuses on their role in creating private wealth (Harhoff *et al.*, 1999) while at the same time generating social benefits (Trajtenberg, 1990b), but more fundamentally in the way in which they hold the potential to transform the economic landscape (Christensen, 1997). Helpman (1998) and Lipsey *et al.* (2005) argue that these transformative powers reside in the broad applicability of many breakthrough innovations that they characterize as general purpose technologies (GPTs). Although the history of development of some GPTs is well-known (e.g. Fogel, 1964; Fishlow, 1965), isolating the introduction of a GPT to the economy and studying its subsequent adoption across the economy has proven difficult (Phene *et al.*, 2006; Kerr, 2010). Feldman and Yoon (2012) argue that the Cohen–Boyer class of patents provides an example of a GPT. To date, the factors that influenced the patterns of adoption of this breakthrough technology remain largely unexplored.

How does one explain the inconstant geography and history of technological advance? A starting point is acknowledging the difficulty of knowledge adoption. Ideas and knowledge are complex goods and Edison's aphorism aside, a precise recipe for their production is unknown. However, with the advent of intellectual property rights protection, knowledge production has become increasingly commodified (Lamoreaux and Sokoloff, 1996), and a critical dimension of competition (Lichtenberg and Philipson, 2002). Nonetheless, the risk of working with new and uncertain knowledge and the attendant high cost of cultivating absorptive capacity cannot be borne by all firms. The search for new technology is highly specialized reflecting the resources and knowledge capabilities of individual economic agents and their partners (Wenerfelt, 1984; Barney, 1991; Kogut and Zander, 1992), the maturity of the industries within which they compete (Abernathy and Utterback, 1978; Klepper, 1997), and the broader ecology of the places where they are located (Cooke *et al.*, 1997; Morgan, 1997; Storper, 1997; Gertler, 2003; Asheim and Gertler, 2005).

Spatial variations in the creation of knowledge and competitive advantage are well known (Feldman, 1994; Maskell and Malmberg, 1999, Feldman and Kogler, 2010). This heterogeneity reflects the pool of private assets and capabilities created by distinct assemblages of firms, workers and institutions in different locations, and by the capacity of these assemblages to develop localized forms of social capital (SAXENIAN, 1994; STORPER, 1997; FELDMAN and ZOLLER, 2012). In relatively thin geographical extensions of these claims the region is viewed as little more than the spatial analogue of the strategic firm partnership. More robust geographical models examine the ways in which spatial proximity increases the flow of tacit knowledge directly through face-to-face contact (MALMBERG and MASKELL, 2002; ASHEIM and GERTLER, 2005), and indirectly through enhancing other forms of proximity within localized clusters of economic actors (GERTLER, 2003).

Arguments about spatial proximity have long played a role in the diffusion of knowledge within geography (HÄGERSTRAND, 1953; BROWN, 1981) and beyond (GRILICHES, 1957). Empirical evidence of the localization of knowledge flows by JAFFE *et al.* (1993), MAURSETH and VERSPAGEN (2002) and SONN and STORPER (2008) reinforce those earlier claims. At the same time, growing recognition of different forms of proximity and relatedness (NOOTEBOOM, 2000; BOSCHMA, 2005; BOSCHMA and FRENKEN, 2010) has raised questions about the role of distance in regulating both the creation and the flow of knowledge. Attention is increasingly directed at the role of social proximity and cognitive proximity in the use and adoption of knowledge (HUBER, 2012).

Social proximity refers to the strength of interpersonal relationships that exist between individual actors (BOSCHMA, 2005). These relationships may take a variety of forms, though they tend to cohere around the concept of trust borne by repeated interaction in common workplaces, industrial organizations or related institutions. AUTANT-BERNARD *et al.* (2007) also note that social proximity can be developed among actors well beyond the local scale often through work-related collaboration, regular meetings, through conferences and trade fairs. Once trust-based social relationships are in place, it is much more likely that actors will engage in interactive learning processes and knowledge sharing, guided by an open attitude towards communicative rationality rather than purely market-driven considerations (LUNDVALL, 1992). Social proximity is much more likely to develop when actors are connected through short social chains. Formal collaboration among individuals, as in the case of co-inventorship, or common employment with the same company, adds to the development of such short chains, that in turn enhances the strength of social proximity (BRESCHI and LISSONI, 2009).

Cognitive proximity focuses upon the extent to which different actors, or in aggregate industries and regions, share common knowledge structures. High cognitive proximity implies greater correspondence between knowledge sets, skills, routines and institutions of knowledge creation and sharing and, thus, a higher potential for absorptive capacity (COHEN and LEVINTHAL, 1990; NOOTEBOOM, 2000). Higher levels of cognitive proximity also lead to enhanced collaboration as well as knowledge sharing. In a similar fashion, recombinant models of technological progress rest on the cognitive proximity of technological subsets and of the economic agents that shape their integration (WEITZMAN, 1998; FLEMING and SORENSON, 2001). KOGLER *et al.* (2013) and RIGBY (2013) extend these arguments in an explicitly spatial framework.

These different forms of proximity are finding purchase in a variety of empirical applications. Thus, BRESCHI and LISSONI (2001, 2004) express a good deal of scepticism regarding the measurement of localized knowledge spillovers, suggesting that empirical estimates are unreliable, at least in part, because they do not separate social from spatial proximity. MAGGIONI *et al.* (2007) develop econometric models exploring knowledge production and co-patenting within and across European regions and show that geographical proximity is always more important than social networks measured by participation within the European Union Fifth Framework Program and European Patent Office (EPO) co-patent applications. Using similar data, AUTANT-BERNARD *et al.* (2007) find strong evidence of spatial and social proximity in research and development (R&D) cooperation across Europe. FISCHER *et al.* (2006) examine patent citations across European regions in an extended gravity model, revealing that spatial and cognitive proximity regulate knowledge flows. In the United States, AGRAWAL *et al.* (2008) use the knowledge production function to explore how spatial proximity and social proximity influence access to knowledge. Using patent citations structured by metropolitan statistical area (MSA) and the co-ethnicity of inventors, they show that the two forms of proximity are statistically significant and that they act as substitutes. STRUMSKY and LOBO (2008) report that the agglomeration of inventors is more important than inventor networks in regulating the pace of invention in metropolitan areas. RIGBY and VAN DER WOUDEN (2013) find that cognitive proximity trumps both spatial and sectoral proximity in this regard. HUBER (2012) provides an excellent summary of much of this work and reports a more nuanced set of results regarding the importance of the different measures of proximity operating within the Cambridge technology cluster.

THE COHEN–BOYER rDNA PATENTS AND THE CREATION OF A NEW SUBCLASS

The Cohen–Boyer discovery builds upon a series of prior technological advances in biochemistry and genetics (FELDMAN *et al.*, 2008). The patent was controversial when filed in 1974 and it was subject to three continuations and a six-year delay. Three factors delayed the granting of the patent (FELDMAN and YOON, 2012). First, academic patents were rare at the time and ownership for discoveries under federally funded research was not automatically assigned to universities until passage of the Bayh–Dole Act. Second, it was unclear whether genetically modified organisms could be patented. This question was answered in the affirmative by the Supreme Court in its ruling on the Diamond–Chakrabarty case. Third, rDNA was highly controversial (SMITH HUGHES, 2001). At the Asilomar Conference organized by Nobel Laureate Paul Berg, the scientific community agreed to a voluntary moratorium on rDNA research until its safety could be investigated.

The original Cohen–Boyer patent application claimed both the process of making rDNA and any products that resulted from using that method. When the USPTO initially denied the product claims, Stanford University divided the original patent application into two divisional product applications, one that claimed rDNA products produced in prokaryotic cells and the other that claimed rDNA products produced in eukaryotic cells[1] along with the process patent. The process patent (USPTO patent number 4237224) is the focus of the present study. The Cohen–Boyer patent has an application date of 1979 and a grant year date of 1980.

Upon granting, every patent is placed into one or more distinct technology classes that are designed to reflect the technological characteristics of the underlying knowledge base that they embody. At the time the Cohen–Boyer patent was granted in 1980, subclass 435/69.1 did not exist. On 5 December 1989, the USPTO issued Classification Order Number 1316, which created a new patent class 435/69.1 – Chemistry: Molecular Biology and Microbiology/rDNA technique that included the method of making a protein or polypeptide. When the set of technology codes is revised, as in this example, the USPTO reviews all granted patents and reclassifies those meeting the criteria of the new codes. At this time, the Cohen–Boyer patent, and all existing patents that made similar knowledge claims, were listed as belonging, at least in part, to class 435/69.1. The USPTO reclassification process provides the researcher with a continuously updated and consistent set of all patents that use specific technologies. STRUMSKY *et al.* (2012) review the use of patent technology codes to study technological change and point to their usefulness in tasks that relate to the identification of technological capabilities, the definition of technology spaces or as an indicator of the arrival of technological novelty.

The data used in this analysis are patent records made available through the USPTO. Patents have become an analytic staple for scholars interested in the geography and history of knowledge production (LAMOREAUX and SOKOLOFF, 1996; Ó HUALLACHÁIN, 1999; JAFFE and TRAJTENBERG, 2002; Ó HUALLACHÁIN and LEE, 2011), on the various types of technical knowledge produced as indicated by patent classes (HALL *et al.*, 2001; STRUMSKY *et al.*, 2012), and on the factors that regulate knowledge flow (JAFFE *et al.*, 1993; BRESCHI and LISSONI, 2001; SONN and STORPER, 2008). The popularity of patent data is related to their ready availability and to the wealth of information they provide. At the same time, the disadvantages of patents as overall measures of economic and inventive activity are well known (PAVITT, 1985; GRILICHES, 1990). It is clear that patents do not represent all forms of knowledge production within the economy and that they do not capture all produced knowledge. Patents, however, do provide insights into the organizations actively engaged in inventive activity in technologies, like rDNA, where the protection of intellectual property is important.

This paper focuses on patents that make knowledge claims in USPTO class 435/69.1, regardless of whether 435/69.1 is the primary class or not.[2] In total, there are 8947 patents used in this analysis. All patents in the sample contain at least one inventor residing in a US MSA.[3] Patents are allocated to the metropolitan areas within which the inventors on those patents are located. If a single patent has two inventors located in different MSAs, then both MSAs are regarded as having knowledge of rDNA technology. Inventors located outside the United States or not located within one of the 366 US metropolitan areas are dropped from the data. Inventors responsible for the production of rDNA patents are identified within USPTO data files using the inventor disambiguation routines of Ronald Lai and colleagues at Harvard (LAI *et al.*, 2011).

The start of the study is 1976, the year of the first USPTO patent application in USPTO class 435/69.1. Three patents predate the application of the rDNA patent (number 4237224) in 1978 because their knowledge claims were adjudicated to belong to class 435/69.1 in a process of reclassification. This paper focuses on the year of application rather than on patent grant year to capture the time of invention. Because many patents are not granted for several years after application, the analysis is ended in 2005 to dampen the impact of right censoring in the data.[4]

THE SPATIAL DIFFUSION OF rDNA

Diffusion of rDNA technology can be traced by 'mapping' the distribution of patents in technology class 435/69.1 across time and space. Fig. 1 shows the growth of rDNA knowledge claims over time, recording the annual count of rDNA patents and the number of metropolitan areas where inventors using rDNA technology resided. The number of patents associated with class 435/69.1 increased rapidly through the late 1980s and early 1990s, following the classic 'S'-shaped diffusion curve. The counts of rDNA patents remain level throughout the late 1990s at around 800 applications per year, although some significant fluctuations are visible. The number of patent

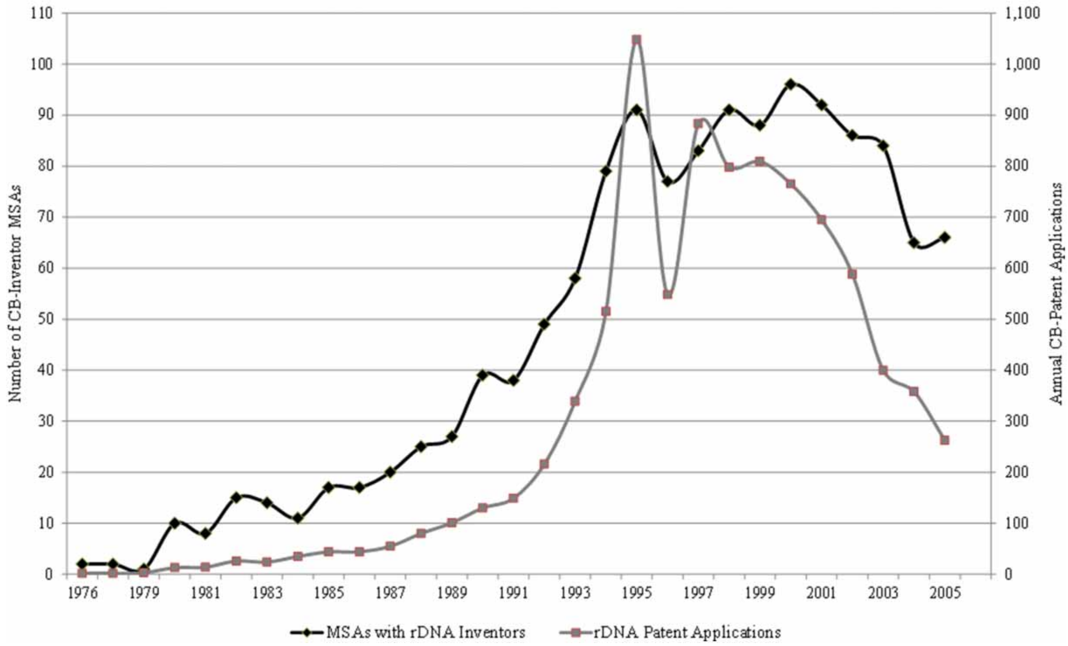

Fig. 1. Annual number of rDNA patent applications and corresponding count of metropolitan statistical areas (MSAs) where their respective inventors reside, 1976–2005

Note: The analysis is based on rDNA patents developed by inventors residing in one of the 366 MSAs) in the United States

applications that make knowledge claims in class 435/69.1 has subsequently levelled off, falling below 300 in 2005, the final year of the investigation.

Few cities were engaged in the production of rDNA inventions before the mid-1980s. By 1987, 11 years after the initial rDNA patent application, only 20 MSAs were producing patents in this technology field. Over the following ten years geographic diffusion accelerated with a little over 90 MSAs developing rDNA technologies in the early 1990s. After stabilizing at this number for about five years, the number of MSAs participating in rDNA invention activities started to decline in 2001. In the final two years observed, 2004 and 2005, the number of MSAs where rDNA invention took place was around 65. Note, however, that right censoring in the data series likely means this number is somewhat higher.

The geographical spread of rDNA technology is further detailed in Table 1 which lists the year of the first rDNA patent in each of the metropolitan areas listed and the year when rDNA-related inventions in each city reached ten patents. The cities listed are well-known centres of invention associated with academic research and subsequently the biotechnology industry. The first ranked city on this list, based on the total number of rDNA-related patent applications from 1976 to 2005, is the metropolitan area around San Francisco (California). Home to Stan Cohen and Herbert Boyer, the two initial inventors of the rDNA technology, this is certainly no surprise. Within San Francisco there is evidence of rapid adoption as a number of different inventors there produced ten patents within three years of the development of rDNA applications. From this initial lead, the city developed a well-known centre of biotechnology research and commercialization activities. The Boston–Cambridge–Quincy (Massachusetts–New Hampshire) metropolitan area, which is also considered one of the key biotechnology centres in the nation, runs a close second to San Francisco, with the same year of initial rDNA patent application but a longer lag of six years to achieve ten patent applications. Total application counts start slightly later in the metropolitan areas of Philadelphia (Pennsylvania), Washington (DC), New York (New York) and San Diego (California).

As awareness of rDNA techniques expanded over space and time, this knowledge subset found broader application as an input to invention across related patent classes. Fig. 2 illustrates the patent classes that have been most frequently combined with the rDNA technology over three year periods running from 1976 to 2005. Technology in class 435/69.1 is most closely associated with its parent class 435: Chemistry: Molecular Biology and Microbiology. While USPC 930: Peptide or Protein Sequence is frequently combined with the rDNA technology in the initial time periods, its significance rapidly declines over time. Moving counter to this, USPC 424: Drug, Bio-Affecting and Body Treating Compositions appears to become increasingly linked to rDNA technology over time, as measured by its co-classification share in patent applications. Over time, the combinations of other

Table 1. Key places (metropolitan statistical areas – MSAs) of rDNA invention

MSA		rDNA patent applications, 1976–2005	Year of first rDNA patent application	Year when MSAs reached ten applications
1	San Francisco–Oakland–Fremont, CA	1133	1978	1981
2	Boston–Cambridge–Quincy, MA–NH	990	1978	1984
3	Philadelphia–Camden–Wilmington, PA–NJ–DE–MD	691	1981	1988
4	Washington–Arlington–Alexandria, DC–VA–MD–WV	639	1980	1986
5	New York–Northern New Jersey–Long Island, NY–NJ–PA	617	1980	1985
6	San Diego–Carlsbad–San Marcos, CA	585	1982	1985
7	San Jose–Sunnyvale–Santa Clara, CA	483	1985	1990
8	Seattle–Tacoma–Bellevue, WA	400	1981	1988
9	Los Angeles–Long Beach–Santa Ana, CA	260	1982	1989
10	St. Louis, MO–IL	150	1976	1989
11	Chicago–Joliet–Naperville, IL–IN–WI	147	1980	1990
12	Sacramento–Arden–Arcade–Roseville, CA	127	1987	1992
13	Baltimore–Towson, MD	126	1988	1993
14	Houston–Sugar Land–Baytown, TX	123	1983	1992
15	Madison, WI	122	1982	1987
16	Indianapolis–Carmel, IN	116	1981	1984
17	Durham–Chapel Hill, NC	113	1984	1992
18	Des Moines–West Des Moines, IA	97	1989	1995
19	Oxnard–Thousand Oaks–Ventura, CA	90	1985	1994
20	Dallas–Fort Worth–Arlington, TX	79	1983	1992

technologies used with USPC 435/69.1 expand. In addition to USPCs 536: Organic Compounds, 530: Chemistry: Natural Resins or Derivatives; Peptides or Proteins, 424: Drug, Bio-Affecting and Body Treating Compositions, 800: Multicellular Living Organisms and Unmodified Parts Thereof and Related Processes, and 514: Drug, Bio-Affecting and Body Treating Compositions, which represent the largest shares of

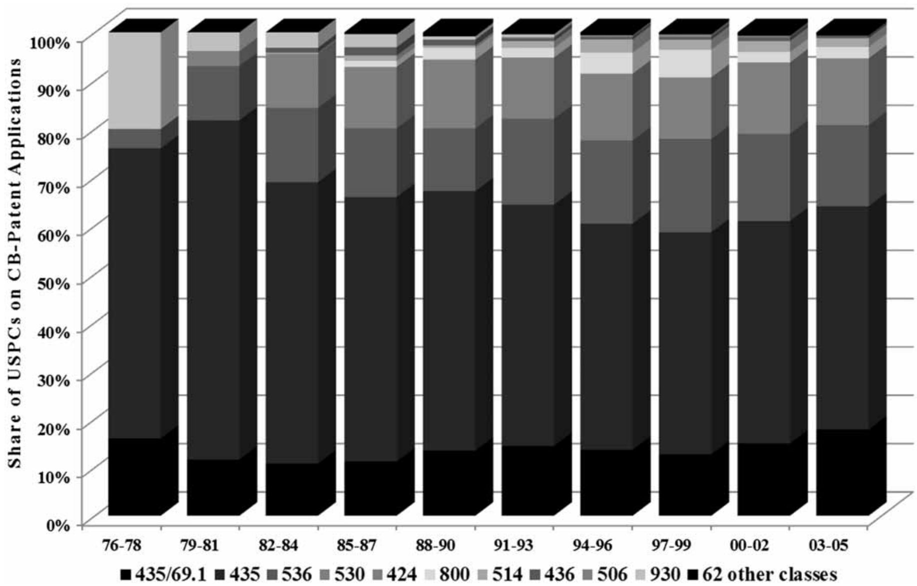

Fig. 2. Distribution of US patent classes (USPC) listed on rDNA-related patent documents; three-year shares based on application year, 1976–2005

Note: 435/69.1 = rDNA USPC 435/69.1, 435 = Chemistry: Molecular Biology and Microbiology, 536 = Organic Compounds, 530 = Chemistry: Natural Resins or Derivatives; Peptides or Proteins, 424 = Drug, Bio-Affecting and Body Treating Compositions, 800 = Multicellular Living Organisms and Unmodified Parts Thereof and Related Processes, 514 = Drug, Bio-Affecting and Body Treating Compositions, 436 = Chemistry: Analytical and Immunological Testing, 506 = Combinatorial Chemistry Technology: Method, Library, Apparatus, and 930 = Peptide or Protein Sequence. The '62 other classes' refers to USPCs that are either rarely combined with class 435/69.1 or to classes that only have been combined with the rDNA technology in more recent time periods, including USPCs 510 (Cleaning Compositions), 977 (Nanotechnology), 426 (Food and Edible Material), and 702 (Data Processing); USPC 514 is an integral part of class 424 combined patent classes, more recently rDNA technology is connected with such diverse technology fields as cleaning compositions (USPC 510), nanotechnology (USPC 977), and data processing (USPC 702).

The primary aim in this paper is to examine the determinants of the spatial diffusion of rDNA technology across US cities. Combining arguments from early research on the diffusion of technology (SILVERBERG, 1991; GEROSKI, 2000; KELLER, 2004) with more recent claims about different forms of proximity (BOSCHMA, 2005) and technological diversification (HAUSMANN and KLINGER, 2007; BOSCHMA *et al.*, 2014; RIGBY 2013) prompts exploration of the following simple model:

$$Y_{ct} = \beta_1 \mathrm{GP}_{ct} + \beta_2 \mathrm{CP}_{ct} + \beta_3 \mathrm{SP}_{ct} + \beta X_{ct}$$

where the dependent variable is binary, indicating whether city (MSA) c develops an rDNA patent in year t; and the three key independent variables represent geographical proximity of a city to rDNA technology (GP), cognitive proximity of a city's knowledge base to rDNA technology (CP) and the social proximity of the city to knowledge about rDNA (SP). The final term in the model captures the influence of a number of covariates that might impact the development of rDNA technology. The interest in these arguments also is related to research examining the relative strength of social proximity and spatial proximity in mediating knowledge flows (JAFFE *et al.*, 1993; BRESCHI and LISSONI, 2001). Analysis of this model takes different forms that are detailed below. First the paper turns to development of the key measures of proximity.

Spatial proximity of rDNA

Invention incorporating rDNA technology depends upon access to knowledge of rDNA. Codified rDNA information may be broadly available, but tacit knowledge of rDNA technology depends upon the ability of a set of potential adopters who have the relevant absorptive capacity. The role of geographical proximity in constraining the flow of rDNA knowledge between US metropolitan areas is operationalized in two ways. First, data on the latitude and longitude of each MSA determine the Euclidean distance between each pair of metropolitan areas. For each city, the average distance to all other 365 metro areas is calculated. A simple hypothesis is that, *ceteris paribus*, metropolitan areas on average closer to all other MSAs are more likely to develop rDNA related capacity in the form of inventions in class 435/69.1. This physical measure of MSA geography is fixed over time. A second measure of geographical proximity combines the distance from city i to all other cities j that have already generated an rDNA patent in class 435/69.1. From this information, two alternative variables capture the changing access of individual cities to the techniques developed by Cohen and Boyer. The first of these variables is the average distance from a city to all other cities that possess knowledge of rDNA technology. The second of the two variables is the distance from each city to its nearest neighbour that has patented in the rDNA technology class. Both variables change over time as the set of MSAs that invent rDNA patents evolves. Theory does not clearly help to differentiate between the alternatives and so both were tried in early empirical work. The minimum distance measure of geographical proximity was much more consistent in the models examined and so the option was for that measure. It is hypothesized that the smaller this minimum distance, the greater the likelihood of a city generating an rDNA patent in a subsequent year. For empirical analysis all independent variables are lagged one year.

Descriptive statistics on this time-varying measure of geographical proximity (minimum distance to rDNA) are reported in Table 2. The minimum distances of zero indicate that a number of cities developed rDNA class patents in the prior year. Those cities that did not develop an rDNA patent in the given year but that are closest to cities that did develop such patents are the top-ranked cities listed at the bottom of Table 2. The maximum value identifies the city that is most remote in terms of having the maximum nearest-neighbour distance to a city that has developed a patent in class 435/69.1. The maximum and mean values of spatial proximity decline over time as expected as the number of cities developing rDNA patents increase through the mid-1990s. Thereafter, these values increase once more, though the mean value of geographical proximity rises only modestly as the set of cities developing rDNA technologies declines late in the study period.

Table 2. Descriptive statistics for city geographical proximity (minimum distance) to knowledge of rDNA

Year	1985	1995	2005
Minimum	0	0	0
Maximum	39.298	6.949	38.825
Mean	3.263	1.111	1.462
SD	3.587	1.190	3.139
Top-ranked cities	Raleigh	State College	Allentown
	Vallejo	Anacortes	Vallejo
	Stockton	Detroit[a]	Salinas
	Providence	Ogden	Daytona Beach[a]
	Ocala	Joplin	Salt Lake City[a]

Note: The distances between cities are Euclidean and based on the geographical coordinates of the centroids of each metropolitan statistical area (MSA). The minimum distances on which the values are based are the distances from each city to its nearest neighbour that developed a patent in class 435/69.1 in the previous year. The top-ranked cities are those without an rDNA patent in the specified year but closest to a city that has. [a]Those cities that developed an rDNA patent the following year.

Social proximity of rDNA

While geographical proximity might represent one measure of relative city access to a certain pool of knowledge, it says nothing about the extent of the interaction that actually occurs between any pair of metropolitan areas. The interaction of particular interest is the potential flow of rDNA-related knowledge. An attempt is made to capture this flow better using information on the spatial distribution of the inventors of rDNA patents and all their collaborators. This measure is operationalized in the following way. First, a city social proximity matrix with dimension 366 × 366 is constructed. All cells in this matrix are coded zero. Second, all inventors of rDNA patents with an application year *t* are identified. A single inventor of an rDNA patent results in a 1 being added to the city social proximity matrix in the cell of the principal diagonal indicating the city where that inventor resides. Co-inventors on a single rDNA patent who live in two cities *i* and *j* result in the value 1 being added to cells (*i*,*j*) and (*j*,*i*) of the city social proximity matrix that is symmetrical. Next, all non-rDNA patents that the rDNA inventors of year *t* have developed over the previous five years are identified. Cross-city collaborations on these non-rDNA inventors result in the value 0.5 being added to the corresponding city-by-city cells of the social proximity matrix. The value 0.5 reflects a discounting of knowledge exchange regarding rDNA that might be present in non-rDNA collaboration. It is to be admitted that this discounting is rather ad hoc and that additional empirical work in this area is warranted. The length of the paper means that such investigation must be reserved for another time. The focus here is on prior co-inventor collaborations for it is anticipated that the knowledge development process takes time and during this development period some knowledge about rDNA might flow to non-rDNA collaborators of rDNA inventors and the former might pass this information to interested parties in their place of residence. The city social proximity matrices are built up annually for the period 1980–2005. Each of these matrices is examined in UCINET and a measure of eigenvector centrality recorded for each city in each year corresponding to the social proximity of the metropolitan area (Borgatti *et al.*, 2002). Note that the social proximity centrality measures are lagged one year in the models below.

Over the first few years after the introduction of the patent for rDNA, the city with the highest social proximity to this technology was Bridgeport (Connecticut), home of Yale University (1980), then Chicago (Illinois) (1981) and then San Francisco (1982). Unsurprisingly, perhaps, San Francisco, the location of the original Cohen–Boyer patent, has the highest measure of social proximity by 1985 and maintains the top-ranked position over the next 20 years. After 1985, larger MSAs, most known for their biotechnology industry clusters, fill out the remaining top ranks of the social proximity measure.

Table 3. Descriptive statistics for city social proximity to rDNA inventors

Year	1985	1995	2005
Minimum	0	0	0
Maximum	50.787	34.786	36.163
Mean	2.214	3.485	3.460
SD	7.063	6.528	6.541
Top-ranked cities	San Francisco	San Francisco	San Francisco
	New York	San Diego	San Jose
	Chicago	New York	San Diego
	Cleveland	Boston	Boston
	Boston	San Jose	Philadelphia

Note: Values are centrality measures from UCINET (Borgatti *et al.*, 2002).

Descriptive statistics on the rDNA metropolitan social proximity measure are reported in Table 3 for three time periods. The relatively high maximum value of social proximity in 1985 indicates the concentration of this technology in patent class 435/69.1, soon after its development. Thereafter, the increase over time in the mean social proximity measure for US metropolitan areas, together with a decline in the maximum value, is evidence of the spatial diffusion of knowledge regarding the Cohen–Boyer technology. The data in Table 3 suggest that between 1995 and 2005 there is little change in the geographical spread of rDNA knowledge.

Cognitive proximity of rDNA

By cognitive or technological knowledge subsets of knowledge are referred to which are associated with particular classes of inventions, technologies or even industries. The proximity of a region to such knowledge subsets refers to local facility or expertise with specific technologies or to how close, in a technological sense, the economic agents of a region are to having such expertise. The knowledge subset of most interest is that circumscribed by patent class 435/69.1. While rDNA patents are a subset of the broader class 435, the subclass is separated in what follows and it is mapped in technology space as a distinct set of knowledge along with 438 other unique primary patent classes.

In order to construct a US knowledge space information is needed on the number of patents in each technology class along with measures of cognitive proximity, the technological distance, between each pair of classes. Co-class information on individual patents is employed to measure the cognitive distance between each pair of technology classes, following the earlier work of Jaffe (1986), Engelsman and van Raan (1994), Verspagen (1997), Breschi *et al.* (2003), and Nesta and Saviotti (2005). The number of primary patent classes of focus here is considerably larger than that employed in most prior studies and thus the technology space outlined below is of higher resolution than those reported to date.

To measure the cognitive proximity, or knowledge relatedness, between patent technology classes in a single year the following method is employed. Let P indicate the total number of patent applications in the chosen year. Then, let $F_{ip} = 1$ if patent record p lists the classification code i, otherwise $F_{ip} = 0$. Note that i represents one of the 438 primary technology classes into which the new knowledge contained in patents is classified, plus an additional class identified as 435/69.1. In a given year, the total number of patents that list technology class i is given by $N_i = \sum_p F_{ip}$. In a similar fashion, the number of individual patents that list the pair of co-classes i and j is identified by the count $N_{ij} = \sum_p F_{ip}F_{jp}$. Repeating this co-class count for all pairs of 439 patent classes yields the (439 × 439) symmetric technology class co-occurrence matrix $\boldsymbol{C}$, the elements of which are the co-class counts N_{ij}. The co-class counts measure the technological proximity of all patent class pairs, but they also are influenced by the number of patents found within each individual patent class N_i. Thus, the elements of the co-occurrence matrix are standardized by the square root of the product of the number of patents in the row and column classes of each element, or:

$$S_{ij} = \frac{N_{ij}}{\sqrt{N_i * N_j}}$$

where S_{ij} is an element of the standardized co-occurrence matrix ($\boldsymbol{S}$) that indicates the technological proximity, or knowledge relatedness, between all pairs of patent classes in a given year. The elements on the principal diagonal of $\boldsymbol{S}$ are set to 1. Alternative forms of standardization are discussed by VAN ECK and WALTMAN (2009) and JOO and KIM (2010).

With the aid of UCINET (BORGATTI *et al.*, 2002), the network of technological relatedness across the 438 primary patent classes and class 435/69.1 is mapped. The technological relatedness 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, and class 435/69.1. The relative positions of the nodes are fixed by the standardized co-occurrence class counts (S_{ij}). Note that the standardized co-occurrence matrix ($\boldsymbol{S}$) is symmetric. The principal diagonal plays no role in the relative locations of the nodes.

The knowledge relatedness networks for 1975–2005 are shown in Fig. 3. The node colours represent the aggregate technology (six class) grouping of HALL *et al.* (2001): black = Chemicals (1), green = Computers and Communications (2), yellow = Drugs and Medical (3), red = Electronics (4), blue = Mechanical (5), and grey = Miscellaneous (6). There is clear evidence of the clustering of individual patent categories within most of these classes, indicating that the technological proximity or relatedness measure is capturing what may be considered as a common knowledge base within these more aggregate technology groupings. 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. rDNA patents in class 435/69.1 are illustrated with the small, yellow triangle in each of the slides of Fig. 3. In early years, rDNA patents are closely linked in the knowledge space with chemicals and with drugs and medical patent classes.

In order to measure the cognitive proximity of the knowledge base of a metropolitan area to the rDNA patent class the average relatedness of a city's patents to class 435/69.1 is found. In technology space, nodes that are close together have a high relatedness score. These are the technology classes that tend to co-occur with relatively high frequency on individual patents. In terms of rDNA, the average relatedness value for metropolitan area m in year t is calculated as:

$$\mathrm{AR}^{mt} = \frac{\sum_j S^t_{\mathrm{CB}j} * D^{mt}_j}{N^{mt}}$$

where S^t_{CBj} represents the technological relatedness between rDNA (class 435/69.1) patents and patents in all other 439 technology classes j, where the vector j includes class 435/69.1. This term is the (row or column) vector of the standardized co-occurrence matrix noted above for the rDNA technology class. D^{mt}_j is the count of the number of patents in technology class j in metro area m in year t; and N^{mt} is a count of the total number of patents in city m in year t.

The measure of cognitive proximity captures the mass of invention that occurs in a particular city and the overall relatedness of that inventive effort to technology class 435/69.1. In general it would be good if the cognitive proximity variable could indicate something about the distribution of a city's inventive efforts in knowledge space along with an index of how close those efforts are to the technology of interest. The chosen term is not ideal, however, for it cannot readily discriminate between small cities that generate few patents that are an average distance away from class 435/69.1 and large cities that generate many patents some close and some far away from class 435/69.1. In this example, whether the small or large city has an advantage in terms of developing rDNA technologies is unclear.

Table 4 presents descriptive statistics for metropolitan area cognitive proximity to the rDNA patent class. The mean average relatedness value of patents in general to the rDNA class of patents within US metropolitan areas was approximately three times higher in 2005 than in 1985, indicating broader use of rDNA and closely related technologies. The metro areas with the highest cognitive proximity values to patent class 435/69.1 are, perhaps, not those that might have been

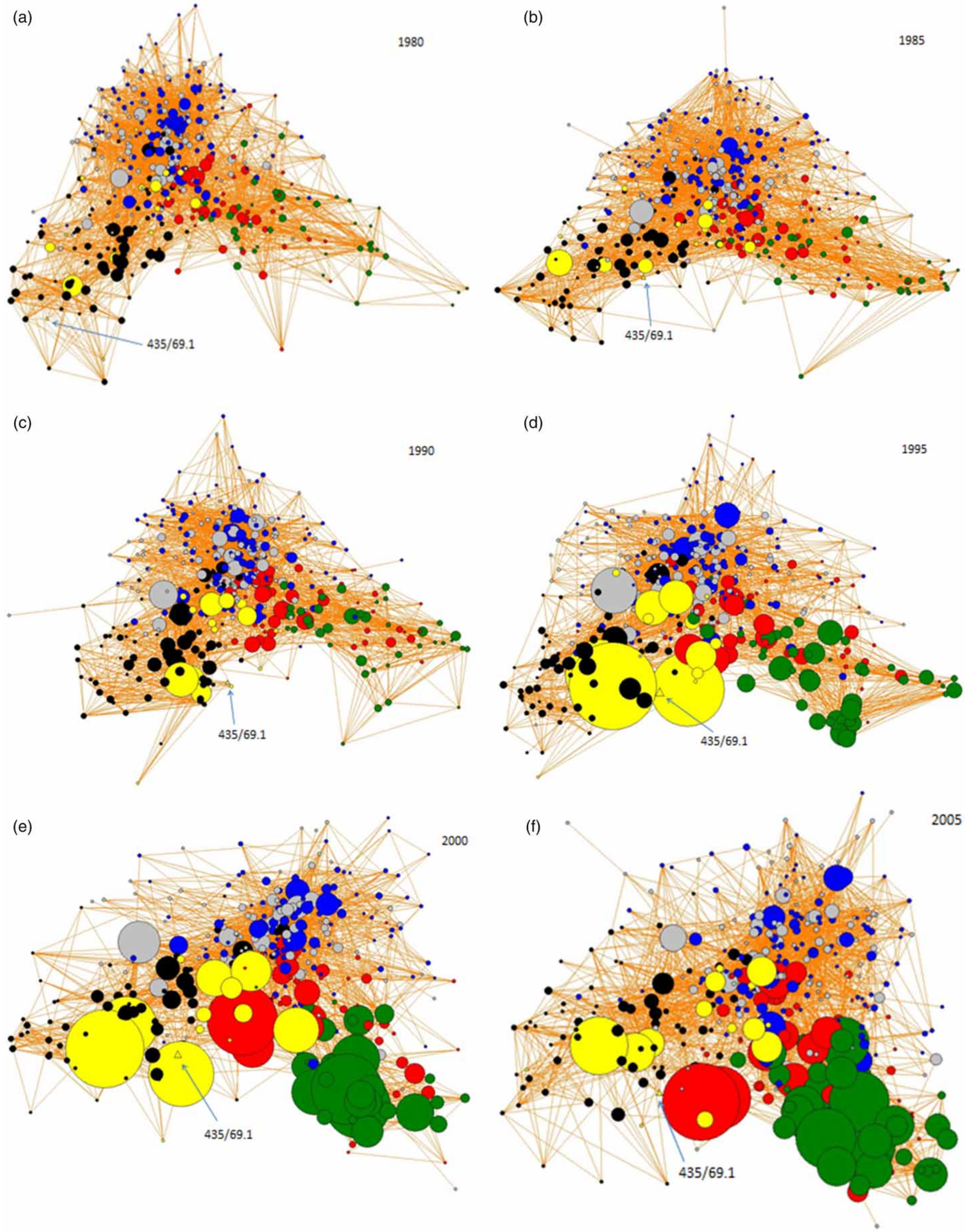

Fig. 3. The US technology space incorporating rDNA (USPC 435/69.1)

Note: Patent class 435/69.1 is the yellow triangle in the lower left of the technology space. The nodes represent all 438 primary classes of utility patents and node sizes reflect the number of patents in each class, scaled for comparability over the years 1980, 1985, 1990, 1995, 2000 and 2005. The colours of the nodes represent the six aggregate technology classes of Hall *et al.* (2001): black = Chemicals (1), green = Computers and Communications (2), yellow = Drugs and Medical (3), red = Electronics (4), blue = Mechanical (5), and grey = Miscellaneous (6)

Table 4. Descriptive statistics for cognitive proximity of metropolitan areas to rDNA

Year	1985	1995	2005
Minimum	0	0	0
Maximum	0.0241	0.0449	0.1160
Mean	0.0016	0.0055	0.0046
SD	0.0029	0.0083	0.0094
Top-ranked cities	Madison	Honolulu	Flagler
	Kennewick	Shreveport	Athens
	Elkhart	Durham–Chapel Hill	Auburn
	College Station	Madison	Iowa City
	Charleston	Blacksburg	Decatur

expected. Most metro areas listed in Table 4 have relatively small numbers of patents, but those patents are in patent class 435/69.1 or close to it in the technology space of Fig. 3. Indeed, Madison (Wisconsin), Kennewick (Washington), Durham–Chapel Hill (North Carolina), Blacksburg (Virginia), Flagler (Florida), Athens (Georgia), and Iowa City (Iowa) are all university towns and sites of rDNA inventions over the period investigated. All the metropolitan areas listed in Table 1 as key centres of rDNA invention have average cognitive proximity values that are greater than average for US cities. It is hypothesized that metropolitan areas with higher levels of cognitive proximity to Cohen–Boyer technology are more likely to develop patents that make knowledge claims in class 435/69.1.

MODEL AND ESTIMATION RESULTS

The primary research question focuses on the probability of a metropolitan area generating an rDNA patent in class 435/69.1. There are time-series panel data available for 366 MSAs over 26 years (1980–2005). The limited (binary) nature of the dependent variable suggests use of a logit or probit regression model. There is also a right-censoring issue in the data that may generate significant bias in estimated coefficients (ALLISON, 1984). Armed with repeated observations on the same set of metropolitan areas over time enables exploration of a fixed effects panel model to deal with potential problems of unobserved heterogeneity. Another possibility that does not control for unobserved heterogeneity, but that more explicitly handles censored data, is the event history model. Both possibilities are explored below using the Cox non-proportional (extended) hazard model, incorporating time-varying covariates, to examine the date of a first rDNA invention within a metropolitan area, while the panel form of the logit model us used to examine the probability of repeated invention in patent class 435/69.1 across all years in the study period.

Time-varying independent variables are lagged by one year in all models employed to generate the results discussed below. There is no clear theoretical rationale for employing a lag of only one period, the aim is only to ensure that the spatial distribution of rDNA patents in year *t* do not influence the value of independent variables constructed for the same year. While one-period lags might dampen some concerns with endogeneity, this is a more serious issue that is returned to below. Attention is first turned to the event history model and at attempts to identify the date at which a metropolitan area first develops a patent in class 435/69.1. The patent data by city are not left-censored for the data series start with the introduction of the first rDNA patent in 1980. However, there are right-censoring issues with the data, as a number of metropolitan areas, 165 out of 366, do not develop a Cohen–Boyer invention by 2005 when the study period ends. In standard regression models that incorporate cross-sectional and time-series data, cross-sectional units that exhibit no variance over time in the dependent variable, are omitted from the model. However, such observations may provide valuable information. For example, a city might never develop a Cohen–Boyer patent because it is geographically remote and inventors in that city might have weak social connections with Cohen–Boyer inventors elsewhere. The authors sought to keep this information in the data set by employing event history techniques. The Cox semi-parametric survival model is the most widely used of the family of hazard models, largely because it does not assume a particular form of probability distribution for survival times. The cost of this flexibility is the assumption of the proportionality of hazards, an assumption that is violated because of the time-varying covariates that enter the model. Thus, the extended Cox model is made use of (BLOSSFELD *et al.*, 2007).

In all the statistical analysis reported below, the paper controls for a number of other covariates that likely influence the spatial diffusion of rDNA technology in addition to the influence of geographical, social and cognitive proximity. The number of patents generated in each metropolitan area provides a proxy for city size/inventiveness. Insofar as patenting in a specialized field of biotechnology is likely associated with basic research in universities and hospitals, typically though not always found in larger urban areas, it is hypothesized that patent counts in general will be positively related to the probability of a city patenting in class 435/69.1. Note that the city-size variable is positively correlated with the social proximity variable, as might be expected. However, that correlation is not cause for undue concern as collinearity renders estimators inefficient rather than biased. Levels of biomedical research funding in universities and in industry were also constructed from National Institute of Health (NIH) awards for each city across the period under study. Higher levels of biomedical research are expected to increase the probability of patenting in rDNA. These

Table 5. Estimating the influence of different forms of proximity on the likelihood of a city inventing a first Cohen–Boyer patent in relation to the baseline hazard (single failure estimated with the extended Cox semi-parametric hazard model with time-varying covariates)

	Hazard ratios		
Time-fixed covariates	Model 1: Full sample	Model 2: Low-patent cities (< 10 patents)	Model 3: High-patent cities (> 90 patents)
Average Distance to Other Cities	1.03451***	1.01377***	1.02711***
	(0.0107)	(0.0484)	(0.0248)
Time-varying covariates			
Lag Geographic Proximity	0.99059	1.05948	0.99897
	(0.0146)	(0.0456)	(0.0243)
Lag Social Proximity	1.03280***	1.29740***	1.00837
	(0.0095)	(0.0916)	(0.0099)
Lag Cognitive Proximity	1.01147***	1.00282	1.04668*
	(0.00243)	(0.0048)	(0.0269)
Lag Patent Count	1.00157***	1.02773	1.00096***
	(0.0002)	(0.0186)	(0.0002)
Lag University R&D	0.99999**	0.99999	0.99999*
	(2.47E–07)	(8.18E–07)	(4.03E–07)
Lag Industry R&D	1.00000***	1.00000	1.00000
	(2.31E–06)	(9.38E–06)	(5.93E–06)
	n = 6573	n = 2397	n = 740
	Failures = 201	Failures = 17	Failures = 87
	LL = –1048.071	LL = –65.944	LL = –292.007
	Prob > Chi^2 = 0.000	Prob > Chi^2 = 0.000	Prob > Chi^2 = 0.000

Notes: All time-varying covariates are lagged one period and are interacted with log(time). The Breslow method is used for ties. Robust standard errors reported in parentheses.

***Significant at the 0.01 level, **significant at the 0.05 level and *significant at the 0.1 level.

LL, log pseudo-likelihood and the overall significance of the model is given by the p-value in the Chi^2 test shown at the bottom of the table. Converting the hazard ratios to a regression coefficient by logging and then dividing by the standard error yields the usual p-scores.

covariates also are lagged one year in all estimations. Finally, also identified in the data were those inventors of rDNA patents who moved cities through the period of investigation and thus took knowledge of Cohen–Boyer technology with them. In only ten (of 201) cases was an inventor move associated with the first year of rDNA invention across all US cities. Omitting these observations in the hazard model resulted in no significant differences in the results discussed below.

Table 5 presents estimation results of the extended Cox hazard model for the patent data against the time-varying covariates. The dependent variable in this model is the number of years after 1980 when a city introduces its first rDNA patent, the hazard. The dependent variable is binary with values of zero representing years prior to the first rDNA patent within a city. The base model is first run and then a series of questions in extensions of the initial specification are examined. Model 1 presents results for all 366 metropolitan areas over all 26 years of the analysis. Two measures of geographical proximity are employed in this first model: the time-fixed measure of the average distance of each city to all its 365 neighbours and the time-varying measure of spatial proximity (minimum distance to rDNA knowledge) defined above. The hazard ratio for the time-fixed measure of the average distance of a city to its neighbours in model 1 is greater than 1 and significant. This means that across all years, MSAs that are on average further from one another have a greater likelihood of developing an rDNA patent. This finding is somewhat unexpected and runs counter to the usual assumptions that information flows more readily over shorter distances rather than long. Turning to the time-varying measure of geographical proximity, the distance of a city to its nearest neighbour that has already developed a patent in class 435/69.1, the hazard ratio is less than 1, consistent with expectations, though it is statistically insignificant. The hazard ratios for social proximity and cognitive proximity are both in line with hypothesized values and are statistically significant. A one-unit increase in social proximity raises the probability that a metropolitan area will generate a first rDNA patent over the baseline hazard by 3.3%. Cognitive proximity raises the hazard ratio by a little more than 1% for every one-unit increase in this variable. Thus, cities that have strong social ties to inventors who have knowledge of rDNA technologies, and cities that have a knowledge base that is relatively close to rDNA technologies are more likely to develop rDNA patents than cities that do not have such characteristics. Examining the other covariates in model 1, large cities, those that patent more in general, tend to develop rDNA patents before smaller cities, and this effect is significant. Similarly, cities with more industrial R&D in the biomedical area develop rDNA patents earlier, though cities with more university R&D are less likely to develop rDNA technologies than others. Even though the effects of R&D spending in the biomedical area have a small influence on the hazard ratio, the negative impact of university R&D spending is puzzling.

Note that metropolitan patent counts are highly correlated with social proximity (Pearson coefficient = 0.6) as might be expected. As the number of patents increase within a metropolitan area, the social proximity of the city also increases. Removing the patent count variable from the Cox model doubles the size of the hazard ratio for social proximity, while leaving all other covariates essentially unchanged.

Models 2 and 3 in Table 5 provide results from the extended Cox model for small cities with relatively few patents and for large cities with relatively high patent counts, respectively. These two groups fall just inside the interquartile range of the distribution of metropolitan areas by patent count, corresponding to the 30th and 70th percentiles. (Trying to estimate the model for the lower quartile generated a very small number of failures (patents in class 435/69.1) and no model convergence.) Cities in the bottom quartile of metro areas by patent count generated only 17 of the 201 total first-time patents examined in the Cox model. Cities in the top quartile were responsible for 87 Cohen–Boyer patents. The key differences between these two subsamples are found in the values of the hazard ratios for social and cognitive proximity. On the one hand, for small cities, the hazard ratio for social proximity is very large, indicating that a one-unit increase in the connectedness of the city's inventors to inventors of rDNA patents raises the probability of patenting in class 435/69.1 by about 30% over the baseline rate. Cognitive proximity has no significant influence on the hazard ratio in small cities. On the other hand, in large cities, social proximity has no significant influence on the hazard ratio, while cognitive proximity exerts a significant, positive effect. It is suspected that cities over a certain size threshold have a sufficient level of social proximity to generate rDNA technologies and that further increases in social proximity make little difference to the probability of such events. In large cities, whether the knowledge base is closely related to rDNA technology is a much more important predictor of increases in the hazard ratio.

The unexpected results on the measures of geographical proximity in Table 5 prompted additional exploration of the spread of rDNA technology across US metropolitan areas. The early years of diffusion seem to dominate the results in Table 5 when rDNA technology jumped between larger cities such as San Francisco, Boston, Philadelphia, Washington, New York, St. Louis (Missouri) and San Diego. This appears to be a process more reflective of hierarchical diffusion rather than the standard epidemic model. Only later in the study period is there more growth of rDNA invention in smaller cities that are somewhat closer to one another than the large cities that dominate the early years. To examine this argument further, the extended Cox model is rerun for two different time periods, before and after 1994. The results are displayed in Table 6.

Table 6. Estimating the Cox extended hazard model for different time periods

	Hazard ratios	
Time-fixed covariates	Model 4: For years before 1995	Model 5: For 1995 and later years
Average Distance to Other Cities	1.04523*** (0.0110)	0.99659 (0.0209)
Time-varying covariates		
Lag Geographic Proximity	0.99338 (0.0129)	0.91656* (0.0450)
Lag Social Proximity	1.02815*** (0.0105)	1.05739*** (0.0223)
Lag Cognitive Proximity	1.02582*** (0.0062)	1.00896*** (0.0028)
Lag Patent Count	1.00166*** (0.0002)	1.00160*** (0.0003)
Lag University R&D	0.99999 (2.95E–07)	0.99999 (5.50E–07)
Lag Industry R&D	1.00000 (3.50E–06)	1.00000 (3.65E–06)
	n = 4398	n = 2175
	Failures = 127	Failures = 74
	LL = –670.296	LL = –369.168
	Prob > Chi^2 = 0.000	Prob > Chi^2 = 0.000

Notes: All time-varying covariates are lagged one period and are interacted with log(time). The Breslow method is used for ties. Robust standard errors reported in parentheses.

***Significant at the 0.01 level, **significant at the 0.05 level and *significant at the 0.1 level.

LL, log pseudo-likelihood and the overall significance of the model is given by the p-value in the Chi^2 test shown at the bottom of the table. Converting the hazard ratios to a regression coefficient by logging and then dividing by the standard error yields the usual p-scores.

Table 6 confirms the intuition. Model 4 is restricted to the years up to 1994, while Model 5 runs over the years since 1994. Note that the significance of the time-fixed measure of geographical proximity disappears in the second time-period while the time-varying measure of geographical proximity becomes significant with the expected negative sign (hazard ratio < 1). Thus, after 1994 (see Model 5), a one-unit increase in the distance of cities to their nearest neighbours with knowledge of rDNA is associated with an 8.5% reduction of the baseline hazard ratio, the likelihood of developing an rDNA patent. Since the mid-1990s, the historical geography of rDNA patenting appears to have shifted, leading to a pattern of geographical infilling, with production of patents in

Table 7. Estimating the influence of different forms of proximity on the probability of a city inventing an rDNA patent (repeated events estimated with fixed effects panel models)

Independent variables	Model 1: Fixed effects panel logit log odds	Model 2: Linear probability model instrumental variables (GMM2S)
Lag Geog Proximity	0.00974	–0.13351***
	(0.0169)	(0.0451)
Lag Social Proximity	0.04570***	0.14152***
	(0.0122)	(0.0204)
Lag Cognitive Proximity	0.00767**	0.14572***
	(3.1092)	(0.0271)
Lag Patent Count	0.00032	–0.01033
	(0.0004)	(0.0172)
Lag University R&D	−1.17E–06**	–0.01507**
	(4.93E–07)	(0.0068)
Lag Industry R&D	0.00001**	0.02217***
	(6.47E–06)	(0.0086)
	n = 4975	n = 8418
	Log likelihood = −1207.355	R^2 = 0.0972
	Prob > Chi^2 = 0.000	Root MSE = 0.2731
Under-identification test		
Kleibergen–Paap LM-statistic (Chi^2 p-value)		173.929 (0.000)
Weak Identification test		
Cragg–Donald Wald F-statistic		27.827
Kleibergen–Paap Wald F-statistic		19.566
Over-identification test		
Hansen J-statistic (Chi^2 p-value)		8.834 (0.1831)

Note: All independent variables are lagged one period. ***Significant at the 0.01 level, **significant at the 0.05 level and *significant at the 0.1 level. Year fixed effects included but not reported. A total of 167 cities (4175 observations) are dropped by the conditional logit (Model 1) because of no change in the dependent variable. In Model 2 all independent variables are log transformed and assumed to be endogenous. The instrumental variables used in this model are two-, three- and four-year time-lags of the full set of independent variables. Model 2 is estimated using generalized methods of moments techniques with Newey–West standard errors that are robust to arbitrary serial autocorrelation and heteroskedasticity. Standard errors are shown in parentheses. In tests of under-identification and weak identification.

technology class 435/69.1 increasingly found in cities that are closer together than in the period before 1994. While the roles of social and cognitive proximity along with city size remain similar over the whole time period, it is noteworthy that the impact of university R&D spending on rDNA patent production is insignificant over both time periods in Table 6.

Table 7 shifts the analysis to a fixed effects logit model (Model 1) that allows one to control for unobserved heterogeneity, though at the expense of concerns regarding the right-censoring in the data. The key finding is that the results are broadly consistent with those already reported for the event history model. One marked difference between the Cox extended hazard model and the logistic model is that the probability of repeated rDNA patenting over time is examined in the latter model, while in the former the focus was only on the time to the first rDNA patent. The fixed-time measure of average distance between cities also drops out of the fixed effects model. In the fixed effects logit model of Table 7 is included time fixed effects, though they are not reported. Note also that the conditional form of the fixed effects logit model eliminates 167 cities from analysis because the value of the dependent variable in these cities is unchanged. In almost all these cases the metropolitan areas in question never develop an rDNA patent. (This is the obvious cost of right-censoring.)

The partial logistic regression coefficients reported in Model 1 of Table 7 are log odds ratios, reporting how a one-unit increase in the independent variable influences a change in the log odds of the dependent variable. For all years, the lagged value of geographic proximity, distance to the nearest city that has generated a Cohen–Boyer patent, has no significant influence on the log odds of a patent in class 435/69.1. Social proximity and cognitive proximity have a significant effect on the log odds ratio and both exhibit the anticipated positive sign. For example, a one-unit increase in social proximity raises the log odds of an rDNA patent being invented in a metropolitan area by 0.0457. This is an increase in the odds ratio of a Cohen–Boyer patent of 1.046, after transforming the coefficient. City size, as proxied by the sum of patent counts, has no significant influence on the log odds ratio. Removing the patent count variable yields no change on the social proximity measure in this model. Research and development in the university and in industry significantly influence the log odds ratio, though again in different directions. Industry R&D increases those odds, while university R&D reduces the log odds of an rDNA patent, suggesting that further work might consider the technology transfer orientation and operations at different institutions. While the Bayh–Dole Act passed in 1980, it was not until the later 1990s that the majority of research universities had established tech-licensing offices. Attitudes towards technology transfer were even slower to change to encourage active patenting. Note that marginal effects are not reliably produced for the panel form of the fixed effects logit model.

Model 2 in Table 7 returns to some robustness issues. First, this paper deals with potential concerns regarding the distribution of independent variables by logging them all. It then moves to the more difficult problem of endogeneity. With MSA-level data, no obvious instruments are available to exploit and thus this paper resorts to a series of two-, three- and four-period lags of all independent variables to serve as instruments. Experimenting with additional lags had little impact on the results, except for the significance of the university R&D variable that tended to become insignificant with longer lags. Kleibergen–Paap statistics reveal that the instrumental-variables model is identified and that lagged values of the independent variables are not weak instruments. The Hansen *J*-statistic suggests that the instruments (collectively) may be regarded as exogenous. Concerns with persistence in the data are dampened using Newey–West standard errors that are robust to arbitrary serial correlation and heteroskedasticity. Given all these caveats, the results in Model 2 lend general support to the key claims, though there are one or two notable changes. First, and most important perhaps, the geographical proximity variable is now significant and consistent with the theoretical claims. When the distance from a city to its nearest neighbour with knowledge of rDNA technology falls, then the probability that the city will develop a patent in class 435/69.1 increases. Second, the city-size (patent count) effect remains insignificant in the fixed effects models of Table 7, unlike in the hazard models presented above.

REFLECTIVE CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

This paper traces changes in the geography of adoption of a significant new technology, rDNA, as represented by knowledge creation in USPTO patent class 435/69.1. Between 1980 and 2005 multiple rDNA patents were developed in approximately 200 US metropolitan areas (counting the location of all co-inventors on these patents). The geographical spread of rDNA technology proceeded slowly at first hitting an inflection point in 1990 when both the number of rDNA patent applications and the number of inventing cities increased rapidly. As rDNA technology matured, perhaps around the year 2000, the number of patent applications began to slow down and a shakeout occurred in the number of cities with active rDNA inventors.

The primary interest was in the factors that regulated the spatial spread of this new knowledge class and, in particular, the relative roles of geographical proximity, social proximity and cognitive proximity. This paper follows a tradition in the literature that seeks to understand the relative importance of these different forms of proximity and that illustrates the mechanisms by which knowledge is transmitted and new technology put to use. Overall, the results suggest that social connections were most important in the diffusion of rDNA technologies. Social proximity, measured by the network of rDNA co-inventors within the United States, played a positive and significant role in the spread of rDNA technology. Inventors associated with patents in technology class 435/69.1 passed information on this new knowledge set to their co-inventors located in the same city or in different cities across the country. This suggests that co-inventing relationships provided a mechanism for the diffusion of the technology: individuals learn about new techniques through inventive work with others and then adopt these new technologies in their subsequent work.

Adoption of this new technological information was not automatic, however, and also depended on cognitive proximity, the technological profile of knowledge in a city and the closeness of that profile to the knowledge base of rDNA. The specialization of rDNA technology appears to have limited adoption in early years to those few metropolitan areas with strong concentrations of biotechnology-related activity. Cognitive proximity to rDNA technology appears to have been the critical dimension of absorptive capacity in those cities exposed to this new technology through their social networks.

rDNA techniques spread from San Francisco to a number of relatively large metropolitan areas and to a few small cities around the country. At least initially, these areas were relatively far from one another and thus little evidence was found to support the role of spatial proximity in facilitating the flow of rDNA-related knowledge. After the mid-1990s, however, the pattern of spatial diffusion shifted to one that was significantly influenced by geographical proximity to rDNA knowledge. Analysis of repeated rDNA invention in the fixed effects logit model, after instrumenting for endogeneity, showed that social, cognitive and geographical proximity all played a significant role in the spread of this technology across US metropolitan areas.

Extensions of the event history model revealed that in smaller, less inventive US cities where cognitive proximity to rDNA technology had little overall effect, social proximity played the critical role in the adoption of knowledge in patent class 435/69.1. This indicates that technology adoption within a city can be enhanced by attracting a few key individuals who have strong social ties to inventors elsewhere. This finding supports the logic of state and local promotion of programmes to hire eminent scholars and star scientists (Feldman *et al.*, 2014). Conversely, in larger, more inventive cities, where it might be assumed that social proximity is always relatively high, absorptive capacity appears to hinge on cognitive proximity, on familiarity with related technologies.

Industry and university R&D spending had mixed roles in accounting for the spread of rDNA technology. The negative influence of university R&D is inconsistent with the theoretical priors. Perhaps the use of NIH awards to university research is not capturing the basic rDNA analysis that was anticipated. Additional work is clearly required on this question.

The aim in this paper was to extend empirical research in evolutionary economic geography through an exploration of different dimensions of proximity and the way in which those dimensions influence the spatial diffusion of a particular technology. Related measures may be employed to examine patterns of diffusion for other technologies. It seems reasonable to assume that patterns of technology diffusion will vary across industries, regions and time periods and for incremental rather than for radical technological breakthroughs. The findings question the extent to which different forms of proximity are substitutes for one another. It is shown that once an idea is patented it is codified and that, at least initially, it follows a pattern of diffusion predicated on social relationships. Overall, though, the results presented here surely raise more questions than they answer:

- How do different technologies travel?
- How are technologies reshaped as they move from one region to another?
- How do the impacts of new technologies vary over space as they interact with existing knowledge systems.
- What are the characteristics of those new technologies that dramatically shift the trajectories of innovation within and across regions?
- How does the diffusion of technology alter knowledge variety within regions, the fundamental building block, perhaps, of evolution within the space-economy?

Acknowledgements – The authors thank Ron Boschma and two anonymous reviewers for insightful comments. The participants at the 'Evolutionary Economic Geography Special Sessions' at the 2012 (New York) and 2013 (Los Angeles) Annual Meetings of the Association of American Geographers also provided useful suggestions on earlier versions of the paper. The authors of this paper are listed alphabetically.

Funding – Financial support for part of this research, which is gratefully acknowledged, was provided by the Regional Studies Association through the Early Career Grant Scheme [grant number R13455 – D. F. Kogler].

NOTES

1. A prokaryotic cell is one without a contained nucleus. The Prokaryotic patent is US 4468464, issued on 28 August 1984. A eukaryotic cell has a contained nucleus. The Eukaryotic patent is US 4740470, issued on 26 April 1988.
2. When rDNA or Cohen–Boyer patents are referred to, one is explicitly referring to patents that make claims to producing or using knowledge in class 435/69.1.
3. For a detailed list and definitions of MSAs, see Office of Management and Budget (OMB) (2009).
4. The average time-lag between the application year and grant year of patents that make a knowledge claim in USPTO class 435/69.1 is about 2.5 years in the early 1980s, rising to just over 3.5 years by the end of the period studied. The database utilized in this study provides data for USPTO patents granted up to the end of 2010 (Lai *et al.*, 2011), and therefore right-censoring the data in 2005 is considered a conservative approach.

REFERENCES

Abernathy W. J. and Clark K. B. (1985) Innovation: mapping the winds of creative destruction, *Research Policy* **14**, 3–22. doi:10.1016/0048-7333(85)90021-6

Abernathy W. J. and Utterback J. M. (1978) Patterns of industrial innovation, *Technology Review* **80**, 40–47.

Abrams D., Akcigit U. and Popadak J. (2013) *Patent Value and Citations: Creative Destruction or Strategic Disruption?* Penn Institute for Economic Research Working Paper No. 13–065. University of Pennsylvania, Philadelphia, PA.

Agrawal A., Kapur D. and McHale J. (2008) How do spatial and social proximity influence knowledge flows? Evidence from patent data, *Journal of Urban Economics* **64**, 258–269. doi:10.1016/j.jue.2008.01.003

Allison P. (1984) *Event History Analysis: Regression for Longitudinal Event Data*. Sage, Newbury Park, CA.

Asheim B. and Gertler M. S. (2005) The geography of innovation: regional innovation systems, in Fagerberg J., Mowery D. and Nelson R. (Eds) *The Oxford Handbook of Innovation*, pp. 291–317. Oxford University Press, Oxford.

Autant-Bernard C., Mairesse J. and Massard N. (2007) Spatial knowledge diffusion through collaborative networks, *Papers in Regional Science* **86**, 341–350. doi:10.1111/j.1435-5957.2007.00134.x

Barney J. B. (1991) Firm resources and sustained competitive advantage, *Journal of Management* **17**, 99–120. doi:10.1177/014920639101700108

Baum J. and Haveman H. (1997) Love they neighbor: differentiation and agglomeration in the Manhattan hotel industry, 1898–1990, *Administrative Science Quarterly* **42**, 304–338. doi:10.2307/2393922

Bessen J. (2008) The value of U.S. patents by owner and patent characteristics, *Research Policy* **37**, 932–945. doi:10.1016/j.respol.2008.02.005

Blossfeld H.-P., Golsch K. and Rohwer G. (2007) *Event History Analysis with STATA*. Lawrence Erlbaum Associates, Mahwah, NJ.

Borgatti S., Everett M. and Freeman L. (2002) *UCINET for Windows: Software for Social Network Analysis*. Analytic Technologies, Harvard, MA.

Boschma R. A. (2005) Proximity and innovation: a critical assessment, *Regional Studies* **39**, 61–74. doi:10.1080/0034340052000320887

Boschma R. A., Balland P. A. and Kogler D. F. (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*. doi:10.1093/icc/dtu012

Boschma R. A. and Frenken K. (2010) The spatial evolution of innovation networks: a proximity perspective, in Boschma R. and Martin R. (Eds) *The Handbook of Evolutionary Economic Geography*, pp. 120–135. Edward Elgar, Cheltenham.

Boschma R. A. and Martin R. (2010) The aims and scope of evolutionary economic geography, in Boschma R. and Martin R. (Eds) *The Handbook of Evolutionary Economic Geography*, pp. 3–39. Edward Elgar, Cheltenham.

Breschi S. and Lissoni F. (2001) Knowledge spillovers and local innovation systems: a critical survey, *Industrial and Corporate Change* **10**, 975–1005. doi:10.1093/icc/10.4.975

Breschi S. and Lissoni F. (2004) *Knowledge Networks from Patent Data: Methodological Issues and Research Targets*. KITeS Working Paper No. 150. KITeS, Centre for Knowledge, Internationalization and Technology Studies, Università Bocconi, Milan.

Breschi S. and Lissoni F. (2009) Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows, *Journal of Economic Geography* **9**, 439–468. doi:10.1093/jeg/lbp008

Breschi S., Lissoni F. and Malerba F. (2003) Knowledge-relatedness in firm technological diversification, *Research Policy* **32**, 69–87. doi:10.1016/S0048-7333(02)00004-5

Brown L. A. (1981) *Innovation Diffusion*. Methuen, London.

Christensen C. M. (1997) *The Innovator's Dilemma*. Harvard Business School Press, Boston, MA.

Clark K. B. (1985) The interaction of design hierarchies and market concepts in technological evolution, *Research Policy* **14**, 235–251. doi:10.1016/0048-7333(85)90007-1

Cohen W. M. and Levinthal D. A. (1990) Absorptive capacity: a new perspective on learning and innovation, *Administrative Science Quarterly* **35**, 128–152. doi:10.2307/2393553

Cooke P., Uranga M. and Etxebarria G. (1997) Regional innovation systems: institutional and organizational dimensions, *Research Policy* **26**, 475–491. doi:10.1016/S0048-7333(97)00025-5

Cowan R. and Jonard N. (2004) Network structure and the diffusion of knowledge, *Journal of Economic Dynamics and Control* **28**, 1557–1575. doi:10.1016/j.jedc.2003.04.002

Dosi G. (1982) Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technical change, *Research Policy* **11**, 147–162. doi:10.1016/0048-7333(82)90016-6

Engelsman E. C. and van Raan A. F. J. (1994) A patent-based cartography of technology, *Research Policy* **23**, 1–26. doi:10.1016/0048-7333(94)90024-8

Essletzbichler J. and Rigby D. L. (2007) Exploring evolutionary economic geographies, *Journal of Economic Geography* **7**, 549–571. doi:10.1093/jeg/lbm022

Feldman M. P. (1994) *The Geography of Innovation*. Kluwer, Dordrecht.

Feldman M. P., Colaianni A. and Liu C. K. (2008) Chapter D, 17.22: Lessons from the commercialization of the Cohen–Boyer patents: the Stanford University licensing program, in Krattiger A., Mahoney R. T., Nelsen U., Thompson J. A., Bennett A. B., Satyanarayana K., Graff G. D., Fernandez C. and Kowalski S. P. (Eds) *Intellectual Property Management in Health and Agricultural Innovation: A Handbook of Best Practices, Vol. 1*, pp. 1797–1808. MIHR, Oxford; PIPRA, Davis, CA; Oswaldo Cruz Foundation, Brazil; and *bio*Developments – International Institute, Ithaca, NY.

Feldman M. P. and Kogler D. F. (2010) Stylized facts in the geography of innovation, in Hall B. and Rosenberg N. (Eds) *Handbook of the Economics of Innovation*, pp. 381–410. Elsevier, Oxford.

Feldman M. P., Lanahan L. and Lendel I. V. (2014) Experiments in the laboratories of democracy: state scientific capacity building, *Economic Development Quarterly* **28**, 107–131.

Feldman M. P. and Yoon J. W. (2012) An empirical test for general purpose technology: an examination of the Cohen–Boyer rDNA technology, *Industrial and Corporate Change* **21**, 249–275. doi:10.1093/icc/dtr040

Feldman M. P. and Zoller T. D. (2012) Dealmakers in place: social capital connections in regional entrepreneurial economies, *Regional Studies* **46**, 23–37. doi:10.1080/00343404.2011.607808

Fischer M., Scherngell T. and Jansenberger E. (2006) The geography of knowledge spillovers between high-technology firms in Europe: evidence from a spatial interaction modeling perspective, *Geographical Analysis* **38**, 288–309. doi:10.1111/j.1538-4632.2006.00687.x

Fishlow A. (1965) *American Railroads and the Transformation of the Ante-Bellum Economy*. Harvard University Press, Cambridge, MA.

Fleming L. and Sorenson O. (2001) Technology as a complex adaptive system: evidence from patent data, *Research Policy* **30**, 1019–1039. doi:10.1016/S0048-7333(00)00135-9

Fogel R. (1964) *Railroads and American Economic Growth: Essays in Econometric History*. Johns Hopkins University Press, Baltimore, MD.

Gambardella A., Harhoff D. and Verspagen B. (2008) The value of European patents, *European Management Review* **5**, 69–84. doi:10.1057/emr.2008.10

Geroski P. A. (2000) Models of technology diffusion, *Research Policy* **29**, 603–625. doi:10.1016/S0048-7333(99)00092-X

Gertler M. S. (2003) Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there), *Journal of Economic Geography* **3**, 75–99. doi:10.1093/jeg/3.1.75

Glaeser E. (2000) The new economics of urban and regional growth, in Clark G. L., Feldman M. P. and Gertler M. S. (Eds) *The Oxford Handbook of Economic Geography*, pp. 83–93. Oxford University Press, Oxford.

Gower J. C. (1971) A general coefficient of similarity and some of its properties, *Biometrics* **27**, 857–871. doi:10.2307/2528823

Griliches Z. (1957) Hybrid corn: an exploration in the economics of technological change, *Econometrica* **25**, 501–522. doi:10.2307/1905380

Griliches Z. (1990) Patent statistics as economic indicators: a survey, *Journal of Economic Literature* **28**, 1661–1707.

Hägerstrand T. (1953) *Innovation Diffusion as a Spatial Process*. University of Lund, Lund.

Hall B. H., Jaffe A. B. and Trajtenberg M. (2001) *The NBER Patent Citations Data File: Lessons, Insights and Methodology Tools*. National Bureau of Economic Research (NBER) Working Paper No. 8498. NBER, Cambridge, MA.

Hall B. H., Jaffe A. B. and Trajtenberg M. (2005) Market value and patent citations, *Rand Journal of Economics* **36**, 16–38.

Hall B., Thoma G. and Torrisi S. (2007) *The Market Value of Patents and R&D: Evidence from European Firms*. National Bureau of Economic Research (NBER) Working Paper No. 13426. NBER, Cambridge, MA.

Harhoff D., Narin F., Scherer M. and Vopel K. (1999) Citation frequency and the value of patented inventions, *Review of Economics and Statistics* **81**, 511–515. doi:10.1162/003465399558265

Hausmann R. and Klinger B. (2007) *The Structure of the Product Space and the Evolution of Comparative advantage*. CID Working Paper No. 146. Center for International Development (CID), Harvard University, Cambridge, MA.

Helpman E. (1998) Introduction, in Helpman E. (Ed.) *General Purpose Technologies and Economic Growth*, pp. 1–14. MIT Press, Cambridge, MA.

Huber F. (2012) On the role and interrelationship of spatial, social and cognitive proximity: personal knowledge relationships of R&D workers in the Cambridge information technology cluster, *Regional Studies* **46**, 1169–1182. doi:10.1080/00343404.2011.569539

Jaffe A. (1986) Technological opportunity and spillovers of R&D, *American Economic Review* **76**, 984–1001.

Jaffe A. and Trajtenberg M. (Eds) (2002) *Patents, Citations, and Innovations: A Window on the Knowledge Economy*. MIT Press, Cambridge, MA.

Jaffe A., Trajtenberg M. and Henderson R. (1993) Geographic localization of knowledge spillovers as evidenced by patent citations, *Quarterly Journal of Economics* **108**, 577–598. doi:10.2307/2118401

Joo S. H. and Kim Y. (2010) Measuring relatedness between technological fields, *Scientometrics* **83**, 435–454. doi:10.1007/s11192-009-0108-9

Keller W. (2004) International technology diffusion, *Journal of Economic Literature* **42**, 752–782. doi:10.1257/0022051042177685

Kerr W. R. (2010) Breakthrough inventions and migrating clusters of innovation, *Journal of Urban Economics* **67**, 46–60. doi:10.1016/j.jue.2009.09.006

Klepper S. (1997) Industry life cycles, *Industrial and Corporate Change* **6**, 145–182. doi:10.1093/icc/6.1.145

Kogler D. F., Rigby D. L. and Tucker I. (2013) Mapping knowledge space and technological relatedness in U.S. cities, *European Planning Studies* **21**, 1374–1391. doi:10.1080/09654313.2012.755832

Kogut B. and Zander U. (1992) Knowledge of the firm, combinative capabilities, and the replication of technologies, *Organization Science* **3**, 383–397. doi:10.1287/orsc.3.3.383

Lai R. D., Amour A., Yu A., Sun Y. and Fleming L. (2011) *Disambiguation and Co-authorship Networks of the U.S. Patent Inventor Database (1975–2010)*. Working Paper. Harvard University, Cambridge, MA.

Lamoreaux N. and Sokoloff K. (1996) Long-term change in the organization of inventive activity, *Proceedings of the National Academy of Sciences, USA* **93**, 12686–12692. doi:10.1073/pnas.93.23.12686

Lichtenberg F. R. and Philipson T. J. (2002) *The Dual Effects of Intellectual Property Regulations: Within- and Between-Patent Competition in the US Pharmaceuticals Industry*. National Bureau of Economic Research (NBER) Working Paper No. 9303. NBER, Cambridge, MA.

Lipsey R. G., Carlaw K. and Bekar C. T. (2005) *Economic Transformations: General Purpose Technologies and Long Term Economic Growth*. Oxford University Press, Oxford.

Lundvall B. Å. (1992) *National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning*. Pinter, London.

Maggioni M., Nosvelli M. and Uberti T. (2007) Space versus networks in the geography of innovation: a European analysis, *Papers in Regional Science* **86**, 471–493. doi:10.1111/j.1435-5957.2007.00130.x

Malmberg A. and Maskell P. (2002) The elusive concept of localization economies: towards a knowledge-based theory of spatial clustering, *Environment and Planning A* **34**, 429–449. doi:10.1068/a3457

Maskell P. and Malmberg A. (1999) The competitiveness of firms and regions: 'ubiquitification' and the importance of localized learning, *European Urban and Regional Planning Studies* **6**, 9–25. doi:10.1177/096977649900600102

Maurseth P.-B. and Verspagen B. (2002) Knowledge spillovers in Europe: a patent citations analysis, *Scandinavian Journal of Economics* **104**, 531–545. doi:10.1111/1467-9442.00300

Mensch G. (1975) *Das Technologische Patt: Innovation Überwinden die Depression, Frankfurt, Umschau*; English edn (1979) *Stalemate in Technology: Innovations Overcome Depression*. Ballinger, New York, NY.

Morgan K. (1997) The learning region: institutions, innovation and regional renewal, *Regional Studies* **31**, 491–503. doi:10.1080/00343409750132289

Nelson R. R. and Winter S. G. (1982) *An Evolutionary Theory of Economic Change*. Belknap, Cambridge.
Nesta L. and Saviotti P. P. (2005) Coherence of the knowledge base and the firm's innovative performance: evidence from the U.S. pharmaceutical industry, *Journal of Industrial Economics* **53**, 123–142. doi:10.1111/j.0022-1821.2005.00248.x
Nooteboom B. (2000) *Learning and Innovation in Organizations and Economies*. Oxford University Press, Oxford.
Ó Huallacháin B. (1999) Patent places: size matters, *Journal of Regional Science* **39**, 613–636. doi:10.1111/0022-4146.00152
Ó Huallacháin B. and Lee D. (2011) Technological variety and specialization in urban invention, *Regional Studies* **45**, 67–88. doi:10.1080/00343404.2010.486783
Office of Management and Budget (OMB) (2009) *OMB Bulletin No. 10–02: Update of Statistical Area Definitions and Guidance on Their Uses*. OMB, Washington, DC.
Pavitt K. (1985) Patent statistics as indicators of innovative activities: possibilities and problems, *Scientometrics* **7**, 77–99. doi:10.1007/BF02020142
Phene A., Fladmoe-Lindquist K. and Marsh L. (2006) Breakthrough innovations in the U.S. biotechnology industry: the effects of technological space and geographic origin, *Strategic Management Journal* **27**, 369–388. doi:10.1002/smj.522
Rigby D. L. (2013) Technological relatedness and knowledge space: entry and exit of US cities from patent classes, *Regional Studies*. doi:10.1080/00343404.2013.854878
Rigby, D. L. and Essletzbichler, J. (1997) Evolution, process variety, and regional trajectories of technological change, *Economic Geography* **73**, 269–284. doi:10.2307/144484
Rigby D. L. and van der Wouden F. (2013) *Abandoning the Herfindahl: Knowledge Relatedness and the Rate of Invention in U.S. Cities*. Department of Geography, UCLA, Los Angeles, CA. [paper available from the authors upon request]
Sahal D. (1981) *Patterns of Technological Innovation*. Addison Wesley, Reading, MA.
Saxenian A. (1994) *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Harvard University Press, Cambridge, MA.
Schumpeter J. A. (1942) *Capitalism, Socialism and Democracy*. Harper, New York, NY.
Silverberg G. (1991) Adoption and diffusion of technology as a collective evolutionary process, in Nakienovic N. and Grubler A. (Eds) *Diffusion of Technologies and Social Behavior*, pp. 209–229. IIASA/Springer, Berlin.
Smith Hughes S. (2001) Making dollars out of DNA: the first major patent in biotechnology and the commercialization of molecular biology, 1974–1980, *Isis* **92**, 541–575. doi:10.1086/385281
Sonn J. W. and Storper M. (2008) The increasing importance of geographical proximity in knowledge production: an analysis of US patent citations, 1975–1997, *Environment and Planning A* **40**, 1020–1039. doi:10.1068/a3930
Sorenson O. (2003) Social networks and industrial geography, *Journal of Evolutionary Economics* **13**, 513–527. doi:10.1007/s00191-003-0165-9
Sorenson O. and Audia P. G. (2000) The social structure of entrepreneurial activity: footwear production in the United States, 1940–1989, *American Journal of Sociology* **106**, 424–462. doi:10.1086/316962
Storper M. (1997) *The Regional World: Territorial Development in a Global Economy*. Guilford, New York, NY.
Storper M. and Christopherson S. (1987) Flexible specialization and regional agglomerations: the case of the U.S. motion picture industry, *Annals of the Association of American Geographers* **77**, 104–117. doi:10.1111/j.1467-8306.1987.tb00148.x
Strumsky D. and Lobo J. (2008) Metropolitan patenting, inventor agglomeration and social networks: a tale of two effects, *Journal of Urban Economics* **63**, 871–884. doi:10.1016/j.jue.2007.07.005
Strumsky D., Lobo J. and van der Leeuw S. (2012) Using patent technology codes to study technological change, *Economics of Innovation and New Technology* **21**, 267–286. doi:10.1080/10438599.2011.578709
Trajtenberg M. (1990a) A penny for your quotes: patent citations and the value of innovations, *Rand Journal of Economics* **21**, 172–187. doi:10.2307/2555502
Trajtenberg M. (1990b) *Economic Analysis of Product Innovation: The Case of CT Scanners*. Harvard University Press, Cambridge, MA.
Van Eck N. and Waltman L. (2009) How to normalize co-occurrence data? An analysis of some well-known similarity measures, *Journal of the American Society for Information Science and Technology* **60**, 1635–1651. doi:10.1002/asi.21075
Verspagen B. (1997) Measuring intersectoral technology spillovers: estimates from the European and US Patent Office databases, *Economic Systems Research* **9**, 47–65. doi:10.1080/09535319700000004
Weitzman M. (1998) Recombinant growth, *Quarterly Journal of Economics* **113**, 331–360. doi:10.1162/003355398555595
Wenerfelt B. (1984) A resource based view of the firm, *Strategic Management* **5**, 171–180. doi:10.1002/smj.4250050207
Zucker L. G., Darby M. R. and Brewer M. B. (1998) Intellectual human capital and the birth of U.S. biotechnology enterprises, *American Economic Review* **88**, 290–306.