Industrial and Corporate Change, Volume 24, Number 1, pp. 223–250
doi:10.1093/icc/dtu012
Advance Access published May 9, 2014

Relatedness and technological change in cities: the rise and fall of technological knowledge in US metropolitan areas from 1981 to 2010

Ron Boschma*,, Pierre-Alexandre Balland†,§, and Dieter Franz Kogler‡**

This article investigates by means of US Patent and Trademark Office (USPTO) patent data whether technological relatedness at the city level was a crucial driving force behind technological change in 366 US cities from 1981 to 2010. Based on a three-way fixed-effects model, we find that the entry probability of a new technology in a city increases by 30% if the level of relatedness with existing technologies in the city increases by 10%, while the exit probability of an existing technology decreases by 8%.

JEL classification: O33, R11, L65, D83.

1. Introduction

In evolutionary thinking, knowledge production is often depicted as a cumulative, path-dependent, and interactive process (Atkinson and Stiglitz, 1969; Dosi, 1982; Nelson and Winter, 1982). Because of uncertainty, agents draw on knowledge

*Ron Boschma, CIRCLE, Lund University, Sölvegatan 16, S-22100 Lund, Sweden. e-mail: Ron.Boschma@circle.lu.se

**Ron Boschma, Department of Economic Geography– URU, Utrecht University, Heidelberglaan 2, 3508 TC Utrecht, The Netherlands. e-mail: Ron.Boschma@circle.lu.se

†Pierre-Alexandre Balland, Department of Economic Geography - URU, Utrecht University, Heidelberglaan 2, 3508 TC Utrecht, The Netherlands and CIRCLE, Lund University, Solvegatan 16, S-22100 Lund, Sweden. e-mail: p.balland@uu.nl

‡Dieter Franz Kogler, School of Geography, Planning & Environmental Policy, University College Dublin, Newman Building, Belfield, Dublin 4, Ireland. e-mail: dieter.kogler@ucd.ie

§Main author for correspondence.

© The Author 2014. Published by Oxford University Press on behalf of Associazione ICC. All rights reserved.

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

acquired in the past, which not only provides opportunities but also sets limits to what can be learned (Heiner, 1983; Cohen and Levinthal, 1990). This happens at the organizational level, where knowledge not only accumulates within the boundaries of the firm but also at the level of territories, as demonstrated by the cumulative and often persistent nature of technological specialization in countries and cities (Archibugi and Pianta, 1992; Lundvall, 1992; Malmberg and Maskell, 1997; Cantwell and Vertova, 2004, Sonn and Storper, 2008).

More recently, research efforts have been directed toward the process of geographical diversification. Scholars like Hidalgo *et al.* (2007), Hausmann and Klinger (2007), and Hausmann and Hidalgo (2010) have argued that the existing set of capabilities in a country determines which new industries will be feasible, and most likely, to develop in the future. By analyzing dynamics in the export portfolios of countries, Hausmann and Klinger (2007) showed that countries predominantly move into new export products that are related to their current export basket. Neffke (2009) suggests that capabilities may not move with ease also within countries, and therefore regions are considered to possess specific capabilities that define which new industries are more likely to emerge and develop in the future. Studies by Neffke *et al.* (2011), Boschma *et al.* (2013), and Essletzbichler (2013) on the long-term industrial evolution of regions found that a new industry is more likely to enter a region when it is related to other industries already in place, and that an existing industry had a higher probability to exit a region when it was not, or poorly, related to other industries already present in the region. However, these studies neglect some of the important features of cities that may affect diversification, such as urban density and technological specialization, as emphasized in the agglomeration economies literature. Moreover, these studies make the claim that regional diversification is driven by technological relatedness at the city level, but they analyze this process in terms of industrial dynamics (i.e. the rise and fall of industries in regions). However, analyzing related diversification in cities by means of technological dynamics (i.e. the rise and fall of technological knowledge in cities) would make a more direct link between urban diversification and the underlying technological nature of relatedness in cities.

Therefore, instead of focusing on *industrial dynamics in regions*, in this article, we focus on *technological knowledge dynamics in cities*, and we analyze whether the rise and fall of technological knowledge is shaped by the existing knowledge base of cities. We draw on the agglomeration economies and the innovation studies literature (in particular, on studies that discuss the economic performance and technological diversification of urban centres) to explain technological change in cities. Our main claim is that cities are more likely to diversify into new technologies that are related to their existing local set of technologies. In this respect, we not only point out that city characteristics drive the process of diversification but also the overall set of technologies that are present in cities. Based on patent data from the US Patent and Trademark Office (USPTO), we investigate the long-term evolution of the

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

patent technology class portfolios of 366 US cities for the period of 1981–2010. First, we construct a so-called *technology space* in which we measure the degree of relatedness between 438 technologies (main patent classes). Then, we determine the relatedness between new and disappearing technologies and the set of preexisting technologies in cities. Finally, we estimate a three-way fixed-effects (F.E) model by using linear probability ordinary least squares (OLS) regression. The results indicate that technological relatedness at the city level was a crucial driving force behind technological change in US cities over the past 30 years.

The structure of the article is as follows. Section 2 sets out the main theoretical ideas on technological knowledge dynamics on the urban scale. Section 3 describes the data, and Section 4 outlines the methodology. We explain the way relatedness between patent technology classes was defined, and how we assess the impact of technological relatedness at the city level on the rise and fall of patent classes in US cities. Section 5 presents the findings, and the final section provides a discussion and concluding remarks.

2. Technological change and related diversification in cities

Cities are engines of invention and economic growth (Hall, 1998; Bettencourt *et al.*, 2007). In the past decade, scholars have been engaged in research to determine whether Jacobs' or Marshallian externalities affect urban invention rates (Feldman and Audretsch, 1999; Paci and Usai, 1999; Ejermo, 2005; O'Huallachain and Lee, 2010). Concisely, Jacobs' externalities are associated with an urban structure composed of a variety of technologies that spark creativity, enable the cross-fertilization of ideas among sectors, and thus generate more inventions. By contrast, Marshallian externalities are cost-reducing externalities, in which the technological specialization of a place enables the better matching of skilled labor and input–output transactions and more effective learning by means of knowledge spillovers.

Generally, empirical studies report rather inconclusive results concerning the question whether technological specialization or diversity leads to higher invention rates (Beaudry and Schiffauerova, 2009). Paci and Usai (1999) showed in a study on 784 Italian local labor systems for the period of 1978–1995 that patenting activity is enhanced both by industrial specialization and diversity. Autant-Bernard (2001) found that technological specialization promoted patent activity in French regions. Ejermo (2005) found a positive relationship between technological specialization (as proxied by patent similarity) and patent productivity in Swedish labor market regions, while O'hUallachain and Leslie (2007) and Lobo and Strumsky (2008) found a positive relationship between per capita patent rates and patent specialization in US cities. In their study on the invention portfolios and the patenting intensity of US cities, O'hUallachain and Lee (2010) showed that urban invention rates are affected by technological specialization and diversity, and that the most inventive cities have deep specializations in different technologies.

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

While these studies on urban specialization versus diversity have led to valuable insights, they tend to treat the technological or industrial structure of cities as given, as if they remain the same, while in reality, those urban structures change over time. Moreover, most of these studies (the 2005 study of Ejermo being a notable exception) do not fully characterize the underlying knowledge stock in cities, and thus, the nature of association between the technology/industry classes found in cities remains largely unspecified. Frenken *et al.* (2007) and Neffke (2009), among others, have argued that the technological relatedness or coherence between industries in cities is crucial in this respect, as relatedness determines learning potentials between technologies and industries in cities.

Only recently, studies have taken a more dynamic approach on the technological and industrial structures of territories, and have combined that with a relatedness perspective. In the past, studies have shed light on the cumulative and persistent nature of technological specialization in countries (Archibugi and Pianta, 1992; Lundvall, 1992; Cantwell and Vertova, 2004), but recently research efforts are also frequently directed toward the process of territorial diversification (Kogler *et al.*, 2013). Hidalgo *et al.* (2007) and Hausmann and Hidalgo (2010) argue that existing capabilities in countries affect their possibilities to develop new industries, and that these capabilities are not internationally tradable. Hausmann and Klinger (2007) demonstrate that countries tend to expand their export activities by moving into export products that are related to their present export portfolio, and countries with a wide range of related export products have more opportunities to deploy their capabilities into new related export products.

Economic geographers have claimed that the urban or regional scale might be even more important for this process of related diversification (Boschma *et al.*, 2013), as many capabilities do not move easily within countries as well (Neffke, 2009). In this context, Maskell and Malmberg (1999) point to the significance of "localized capabilities," which are associated with a particular local knowledge base and institutional context. Because regions accumulate specific competences, these offer additional learning opportunities for local organizations and lower search costs for new knowledge in similar fields. Consequently, search behavior for new knowledge tends to be myopic and localized, both in cognitive and geographical terms (Lawson, 1999; Maskell and Malmberg, 2007). Such geographically localized learning is embedded in local institutions, such as social conventions that create mutual understanding between local agents and make them interact and learn (Storper, 1995; Gertler, 2003). These "localized capabilities" are regional intangibles assets with a high degree of tacitness that are difficult to replicate in other places.

Only recently, there is a growing awareness that these geographically localized capabilities also operate as a key source of technological diversification (Boschma and Frenken, 2011). Technological diversification is accompanied with high risks and switching costs because the capabilities of firms and their embeddedness in the local environment clearly limit the possibilities to move in completely different technology

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

sectors and markets. Therefore, usually when firms diversify into new technologies and products, they will stay close to their existing capabilities (Penrose, 1959; Teece *et al.*, 1994; Antonelli, 1995; Breschi *et al.*, 2003; Piscitello, 2004) and remain in the same location where they can more easily draw on related capabilities (Frenken and Boschma, 2007; Buenstorf and Guenther, 2011). There is strong evidence from longitudinal studies on industries that many successful entrepreneurs in new industries do exploit regional competences they previously acquired in technologically related industries (Klepper, 2007; Buenstorf and Klepper, 2009), in particular, during the infant stage of the industry (Boschma and Wenting, 2007). It is also likely that new industries recruit skilled labor from local related industries and benefit from that, as the local supply of related skills enables easier matching of labor and enhances learning processes (Eriksson, 2011).

A large body of descriptive studies has demonstrated that new local industries are indeed rooted in related regional activities (Chapman, 1991; Glaeser, 2005; Bathelt *et al.*, 2011). Recently, more quantitative studies (Neffke *et al.*, 2011; Boschma *et al.*, 2013) have focused on this process of related diversification in a large number of regions in countries such as Sweden and Spain over a long period. These studies found systematic evidence that new industries are more likely to enter a region when these are technologically related to other industries in that region. Another interesting finding was that an existing industry has a lower probability to exit a region when that industry is technologically related to other industries in the region. This latter finding is as expected, considering that these industries are more centrally positioned and more fully embedded in the local networks of related industries and because they are better capable of securing their vested interests through their strong ties with other local stakeholders, including policymakers (Hassink, 2010).

In essence, high costs prevent regions to build completely new industries from scratch and to abandon existing industries that are deeply rooted locally. Thus, it is not surprising that empirical studies find that the rise and fall of industries in regions is subject to a path-dependent process, which is driven by the degree of technological relatedness with other local industries. This also explains why the industrial structure in regions is most likely technologically cohesive, something that tends to persist over time despite the fact that industries come and go (Rigby and Essletzbichler, 1997; Neffke *et al.*, 2011; Rigby, 2013; Essletzbichler, 2013). It is not the lack of industrial dynamics, but precisely the actual occurrence of (quite regular patterns of) industrial dynamics that makes the techno-industrial structure of regions rather cohesive. This is mainly due to the exit or loss of existing industries, which tends to lower variety but increase related variety in regions, as more unrelated industries are more likely to disappear. Although the entry of new industries injects new variety into regions, this will not concern just any industry, but rather industries that are technologically related to other regional industries (Neffke, 2009).

The studies on related diversification briefly mentioned above have focused on industrial dynamics in regions, and how the degree of technological relatedness with

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

existing regional industries impacted on the rise and fall of industries. While providing important insights, these studies on regional diversification also suffer from two theoretical shortcomings. First, they have neglected features such as population density and technological specialization of cities that may affect the diversification process, as emphasized in the agglomeration economies literature. Second, although these studies argue that regional diversification is driven by technological relatedness at the regional scale, they analyze this process by means of the rise and fall of industries in regions. However, it would be more plausible to analyze related technological diversification in cities by means of technological dynamics instead of industrial dynamics, as a study on the rise and fall of technological knowledge in cities would establish a more direct link between related urban diversification and its underlying technological nature.

Therefore, this article analyzes whether the rise and fall of inventions, or more precisely, the entry and exit of patent technology classes in cities is conditioned by the existing set of technological knowledge in these metropolitan areas. It is expected that technological relatedness at the urban level is a key driving force behind technological change in cities. In addition, it is also expected that other more general features of cities such as population density and technological specialization might affect the process of technological diversification in cities. To test this, we investigate the evolution of patent portfolios in 366 US cities for the period of 1981–2010.

3. Constructing the dataset

Patents and patent statistics encompass an incredible wealth of information with the potential to facilitate a multitude of approaches in the investigation of knowledge creation and diffusion processes (Scherer, 1984; Griliches, 1990; Jaffe and Trajtenberg, 2002). Patent statistics are considered a "noisy" indicator when utilized as an overall measure of economic and inventive activity, mainly because patented inventions do not represent all forms of knowledge production with an economy and thus certainly do not capture all produced knowledge (Pavitt, 1985; Griliches, 1990). Nevertheless, if the focus is on economic valuable technical knowledge that pertains to inventions of utility, patents provide an excellent opportunity for the study of technological change. For example, more recently, patent data have been utilized to study the evolution of technologies by taking advantage of the largely unexploited information of technology classes that are listed in patent documents (Fleming and Sorenson, 2001; Nesta, 2008; Quatraro, 2010; Strumsky *et al.*, 2012). Following this lead, the aim of our analysis is to extend this approach and empirically test how the presence of, and relatedness among, patent classes shapes technological change in an urban setting. Specifically, the goal is to outline a model that describes the emergence, as well as the exit, of new technologies in US metropolitan areas from 1981 to 2010.

There are several patent databases that are publicly available for research purposes. Two prominent examples are the "Patent and Citations Data" of the National Bureau

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

of Economic Research (Hall *et al.*, 2001) and the "Patent Network Dataverse" from the Institute for Quantitative Social Science at Harvard University (Lai *et al.*, 2011). The USPTO served as original data source for both of these, and further provides supplement information, i.e. "the USPTO Harmonization of Names of Organizations Data File" (USPTO, 2010), that allowed for an extension of these databases, which are utilized in the present study. To facilitate an analysis pertaining to technological change in US cities, individual patent documents were allocated to one of 949 Core-Based Statistical Areas (CBSAs) based on the first inventor's residency record (OMB, 2010). For more recent records, this was an effortless task owing to the availability of ZIP codes. However, for some of the older records in the available patent databases, it was necessary to use the geographical correspondence engine available through the Missouri Census Data Center in order to link inventor records to their respective cities of residency at the time the invention was developed. It was then deemed to be reasonable to only focus on the 366 Metropolitan Statistical Areas (MSAs) rather than the whole population of cities that also includes micropolitan areas, simply because MSAs house well over 90% of all patented utility invention in the United States over the past three decades. In order to produce results that reflect the real timing of inventive activity, and thus the entry and exit of technologies in cities, all indicators that were developed in the data set subsequently are based on the application rather than on the grant dates listed in the original patent documents. This is also mainly to avoid time-skewed results due to the continuously increasing time lag from the date of invention and filing to the grant date over a 30-year time frame.

Patents are classified into one or more distinct technology classes that reflect the scope of the approved claims listed in a patent document. Based on the available data, there are 438 main patent classes, i.e. technology codes that utility patents have been assigned to by the USPTO over the past three decades. It should be noted that this refers to the number of patent-specific codes that were up-to-date by the end of 2010 rather than all codes that the USPTO has ever utilized since it was established more than 200 years ago. In essence, the USPTO redefines classes, adds new ones, and sometimes, although rarely, even abandons existing ones on an ongoing basis. All of this is documented in the monthly "classification orders" that are issued by the USPTO, indicating changes to the definition of the classification system at a given time. The advantage of this continuous process is that it provides a consistent set of technology classes into which patents are placed, something essential in an investigation that relies on data collected over a prolonged period. Strumsky *et al.* (2012) provide a detailed account of patent technology codes and how these should be interpreted. The number of 438 main patent classes utilized in the present analysis also corresponds to the number used in other relevant studies, including Rigby (2013) and Kogler *et al.* (2013). The spatial and technology codes that have been constructed and identified in the patent database briefly outlined here will serve as a point of departure for the analysis that follows.

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

4. Indicators of relatedness and econometric model

As explained in Section 2, we expect relatedness to be a major driving force behind technological change in cities over time. We argue that new technologies emerge from the recombination of existing technological knowledge, leading to the diversification of cities into technological activities that are related to their specific knowledge structure. To test these predictions, we follow the methodology developed and applied in recent studies by Hidalgo *et al.* (2007), Neffke *et al.* (2011), Rigby (2013), and Boschma *et al.* (2013). First, we construct a so-called *technology space* in which we measure the degree of relatedness between all technologies. Second, we determine the relatedness between new and disappearing technologies and the preexisting technological knowledge structure of cities, which we define as relatedness density (density of related technologies). We use USPTO patent data to regress the entry and exit of technological activities in US cities during the period of 1976–2010 on the degree of technological relatedness of these activities with the existing ones in cities.

4.1 The technology space

To measure the relatedness between technologies and draw the resulting *technology space*, we follow the *product space* framework proposed by Hidalgo *et al.* (2007). The *product space* is a network-based representation of the economy, where the nodes define product categories and the ties between them indicate their degree of relatedness. The main idea developed by Hidalgo *et al.* to capture relatedness is to look at how often two products are exported by countries. Two products are then considered to be related if they are co-exported by many countries because they are assumed to require the same capabilities following the reasoning outlined in detail above.

Using this framework, we construct the *technology space*, which is a network-based representation of the relatedness between all the technologies that can be found in the domestic patent portfolio of the United States from 1976 to 2010. In this $n*n$ network, each node i ($i = 1, \ldots, n$) represents a specific technological class. Applying the measure to the three-digit USPTO main patent classes (Hall *et al.*, 2001), we are able to map the degree of relatedness between 438 different technological classes. For instance, one of the biggest nodes in this network represents the technological *class 800* ("*multicellular living organisms*"), which is a subcategory of the biotechnology class[1]. To compute the degree of relatedness between all these 438 technologies and draw the resulting network, we focus on how often two technologies are found within the same US city, defined as an MSA.[2]

[1] Class 33 listed in the study by Hall *et al.* (2001).

[2] Patents were assigned to each MSA according to the residency information provided by the primary inventor of a patent; see Section 3 for further information regarding the spatial allocation of patent documents applied in the present study.

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

The relatedness $\varphi_{i,j,t}$ between each pair of technology i and j is computed by taking the minimum of the pair-wise conditional probabilities of cities patenting in one technological class i, given that they patent in another technological class j during the same period. To avoid the noise induced by negligible patenting activity, we only consider the primary technological classes listed on patent documents in which cities have a revealed comparative advantage (RCA), as proposed by Hidalgo *et al.* (2007).

$$\phi_{i,j,t} = \min\left\{P\left(RCAx_{i,t}|RCAx_{j,t}\right), P\left(RCAx_{j,t}|RCAx_{i,t}\right)\right\} \tag{1}$$

where a city c has an RCA in technology i in time t if the share of technology i in the city's technological portfolio is higher than the share of technology i in the entire US patent portfolio. More formally, $RCA_{c,t}(i) = 1$ if

$$\frac{patents_{c,t}(i)/\sum_i patents_{c,t}(i)}{\sum_c patents_{c,t}(i)/\sum_c \sum_i patents_{c,t}(i)} > 1 \tag{2}$$

We compute the relatedness $\varphi_{i,j,t}$ between each pair of technologies i and j for 7 periods of 5 years, from 1976 to 2010. As presented in Table 1, the correlation between these seven matrices of relatedness is very high, indicating that technological change is a slow, gradual, and path-dependent process. Figure 1 provides a visual impression of the technology space based on the average degree of relatedness for the entire period of 1976–2010. From this graph, it is clear that the different technological classes tend to form interconnected groups that closely correspond to the classification[3] in six main technological areas (mechanical, chemical, drugs and medical, electrical and electronic, computers and communications, and others) as proposed by Hall *et al.* (2001).

As a robustness check, however, we will also verify whether the econometric results hold for alternative measures of relatedness. First, we constructed similar network matrices of the technology space by measuring relatedness through normalized co-occurrences. Following iCancho and Solé (2001) in the context of co-occurrences of words within sentences, we consider that technological classes are related when they co-occur more than one can expect by chance, that is, when the observed co-occurrences are higher than the expected values based on the probability calculus. Second, we constructed a relatedness matrix where technological classes are related if they are listed in the same (two digits) subcategory[4] of the patent classification proposed by Hall *et al.* (2001).

[3] Rigby (2012) also found a strong correlation between the classification of Hall *et al.* (2001) and a relatedness graph constructed from patent citations. This is somewhat expected considering that the search process of prior art and the assignment of technology classes, both of which are carried out by patent examiners, are interrelated tasks; see Alcácer and Gittelman (2006: 778).

[4] Hall *et al.* (2001) proposed 36 patent categories at two-digit levels.

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

Table 1 Change in relatedness between technologies (1976–2010)

Time periods	1976–1980	1981–1985	1986–1990	1991–1995	1996–2000	2001–2005	2006–2010
1976–1980	1.000	–	–	–	–	–	–
1981–1985	0.800	1.000	–	–	–	–	–
1986–1990	0.750	0.798	1.000	–	–	–	–
1991–1995	0.725	0.768	0.836	1.000	–	–	–
1996–2000	0.711	0.740	0.802	0.852	1.000	–	–
2001–2005	0.690	0.720	0.782	0.825	0.868	1.000	–
2006–2010	0.597	0.638	0.688	0.702	0.735	0.773	1.000

4.2 Relatedness density of US cities

To analyze how relatedness influences technological change within cities, we need to construct a city technology–level variable[5] that indicates how close a potential new technology is to the existing technological portfolio of a given city. This idea is operationalized by the density index (Hidalgo *et al.*, 2007), which measures, in our case, the proximity of a new technology to the existing set of technologies in a given city. More formally, the density around a given technology i in the city c in time t is computed from the technological relatedness[6] of technology i to the technologies in which the city c has an RCA in time t, divided by the sum of technological relatedness of technology i to all the other technologies in the United States in time t:

$$RD_{i,c,t} = \frac{\sum_{j \in c, j \neq i} \phi_{ij}}{\sum_{j \neq i} \phi_{ij}} \times 100 \tag{3}$$

By construction, this relatedness density variable lies between 0% and 100%. A city technology–level density equal to 0% indicates that there is no technology related to technology i in the city c, while a value of 100% indicates that all the technologies related to technology i belong to city c's technological portfolio. To take

[5] In the econometric models presented in this article, the unit of observation refers to city–technology pairs, over several periods.

[6] To ease the interpretation of the econometric results, we dichotomize the relatedness index. In the present article, we use a 5% threshold, which means that the top 5% of all technology pairs that have the highest $\varphi_{i,j}$ are considered as related, while the remaining 95% are considered as unrelated. The results presented in this article, however, are robust to numerous other specifications using not only the valued relatedness index but also to the dichotomized index when we change the threshold to 1%, 10%, and 20%.

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

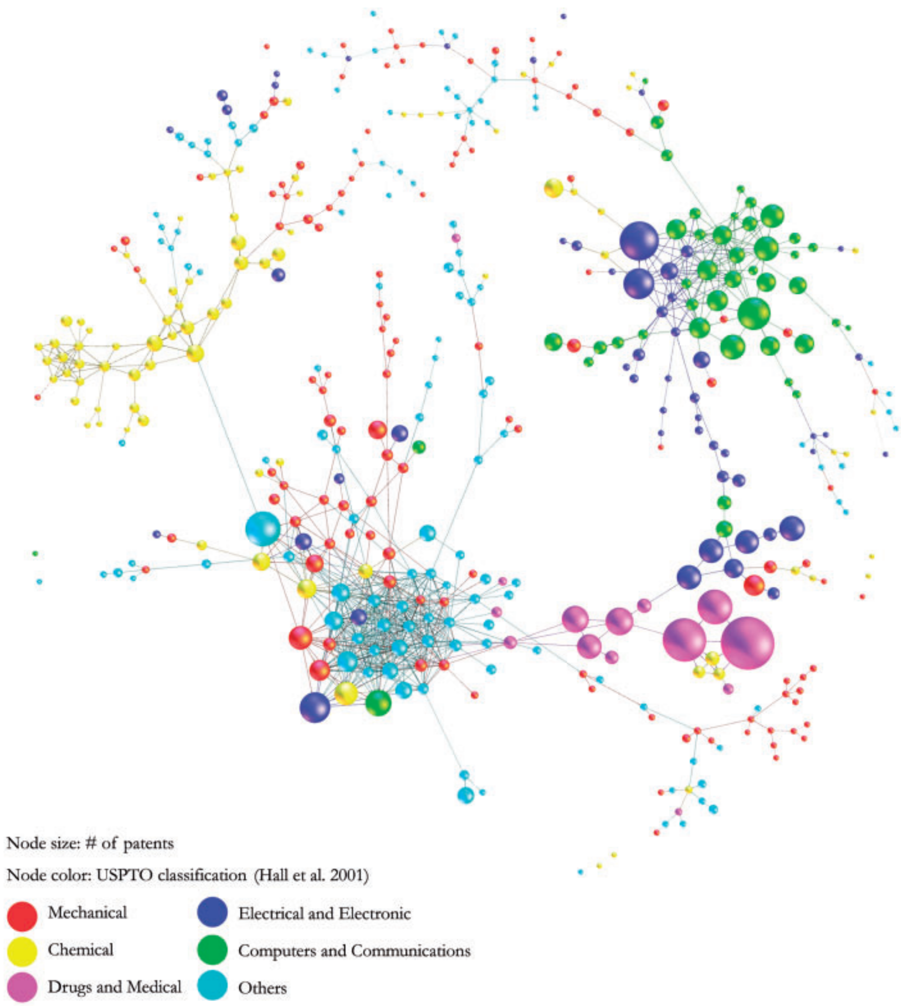

Figure 1 The US technology space (1976–2010). Notes: The "technology space" is constructed in a similar fashion than the "product space" proposed by Hidalgo *et al.* (2007). Each node ($n = 438$) represents a patent technology class (Hall *et al.*, 2001), and the links between these patent classes indicate their technological relatedness.

a concrete example, the density around the technological class "Semiconductor Devices" (class 438) in the Atlanta metropolitan area[7] was equal to 52% during the period of 1981–1985. In fact, the technological class "Semiconductor Devices" was related to 34 other classes in total, and 18 of these 34 classes belonged to Atlanta's technological portfolio at that time. In the next period (1986–2010), class 438 actually emerged in Atlanta, which follows the expectation that the density of related technologies shapes technological change in cities.

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

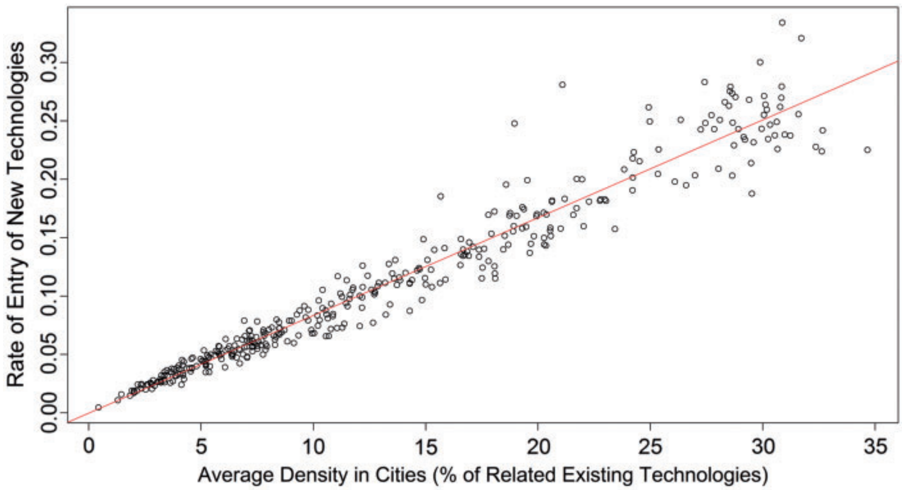

Figure 2 Relatedness and technological change in US cities (1981–2010). Notes: Each dot represents one of the 366 US cities (MSA). The rate of entry represents the number of new technologies that entered a city's technology space divided by the total number of possible entries. Average density is the average percentage of related technologies in the city.

As Figure 2 shows, a strong and positive relationship seems to exist between relatedness density and the emergence of new technologies in cities. To draw Figure 2, we plotted the average density values and the rate of entry of new technologies for each of the 366 cities. The rate of entry is given by the sum of entries of new technologies in a city from 1986 to 2010, divided by the total number of possible entries during this period. For instance, one of the metropolitan areas with the highest rate of entry is the Greater Boston area[7]. On average, at each of the six 5-year periods, 20 new technologies entered, while 80 could have entered. This indicates a probability of entry of about 25%, while the average rate of entry in American cities is about 9%[8]. The average density around these potential technologies was also about 25%. Figure 2 shows that there is a very high positive correlation between the level of relatedness density and the probability of entry. Cities that have a more diverse technological portfolio and have competences in core technologies (i.e. at the core rather than at the periphery of the technological space displayed in Figure 1) seem to renew themselves more quickly over time. But of course, the Greater Boston area is a rich metropolitan area, which hosts very productive research universities and also scores high in terms of human capital (Glaeser, 2005).

[7] Boston–Cambridge–Quincy Metropolitan Statistical Area.

[8] See the mean of entry in Table 2.

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

Therefore, to test the relationship between relatedness and technological change at the city level, one has to control for city and technology time-variant and time-invariant characteristics. This is the purpose of the econometric framework presented further below.

4.3 Econometric specifications

We want to estimate how relatedness influences technological change at the city level. Therefore, we regress the emergence of new technologies on their degree of relatedness with the technological portfolio of cities, which is captured by the relatedness density variable. The basic econometric equation to be estimated can be written as follows:

$$Entry_{i,c,t} = \beta_1 Density_{i,c,t-1} + \beta_2 City_{c,t-1} + \beta_3 Techno_{i,t-1} + \phi_c + \psi_i + \alpha_t + \varepsilon_{i,c,t} \tag{4}$$

where the dependent variable $Entry_{i,c,t} = 1$ if a technology i that did not belong to the technological portfolio of city c in time $t-1$ enters its technological portfolio in time t, and 0 otherwise. The key explanatory variable $Density_{i,c,t-1}$ indicates how the potential new technology i is related to the preexisting technological set of capabilities of city c.

$City_{c,t\text{-}1}$ is a vector that summarizes a range of observable time-varying city characteristics[9]. We constructed variables such as the number of employees in a city (*employment*), the number of inhabitants by square meters (*population density*)[10], the ratio of inventors to employees (*inventive capacity*), the growth rate of the number of inventors (*MSA technological growth rate*), and the economic wealth of a city (*income per employee*)[11]. The variable *technological specialization* of city c has been measured by the average location quotient weighted by the number of patents:

$$Specialization_c = \sum_i \frac{P_{ci}}{P_c} LQ_{ci} \tag{5}$$

[9] We would like to thank one of the referees for pointing out important control variables at the city and technology level.

[10] The original data source for county population data was the NBER (http://www.nber.org/data/census-intercensal-county-population.html). The country land area was obtained from the 2010 Gazetteer File (US Census at http://www.census.gov/geo/maps-data/data/gazetteer.html). All of these data are on the county level. Thus, we had to allocate counties to MSAs subsequently. This was done by means of a table that was obtained from the Bureau of Economic Analysis (http://www.bea.gov/regional/docs/msalist.cfm). After aligning all the county data, it was then possible to generate population density at the MSA level for all the different periods.

[11] The data source for table "CA30 Regional Economic Profiles—Earnings by place of work (thousands of dollars)" was the Bureau of Economic Analysis. In order to get a per capita measure, we applied the table "CA30 Regional Economic Profiles—Wage and salary jobs (number of jobs) from the same agency.

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

where *Pci* denotes the number of patents of city *c* in class *i*, *Pc* the total number of patents of city *c*, and *LQ* the location quotient of technology *i* in city *c*.

$Techno_{i,t-1}$ is a vector that summarizes a range of observable time-varying technological characteristics. First, we take the total number of inventors (*Nb. Inventors*) computed at the technology level to control for technology size, as technologies that involve many inventors are more likely to enter any city by chance. We included two variables to account for technology age and the expansion/extraction of technological opportunities, by using the year in which a technological class has been officially established by the USPTO (*date established*) and the growth rate of the number of inventors patenting in a given technology class (*Tech. Class growth rate*). As concentration of inventive activities could be more conducive to related technological diversification, we also measured the average location quotient weighted by the number of patents (*technological concentration*):

$$Concentration_i = \sum_c \frac{P_{ci}}{P_i} LQ_{ci} \tag{6}$$

where P_{ci} denotes the number of patents of class *i* in city *c*, P_i the total number of patents of class *i*, and *LQ* the location quotient of technology *i* in city *c*.

Finally, ϕ_c is a city-fixed effect, ψ_i is a technology-fixed effect, α_t is a time-fixed effect, and $\varepsilon_{i,c,t}$ is a regression residual.

Therefore, the baseline econometric model was used is a three-way F.E model, to take into account omitted variable bias at the city and technology levels, assuming that these omitted variables are constant over time. We estimate equation (4) by using a linear probability (OLS) regression[12]. ϕ_c, ψ_i, and α_t fixed effects are directly estimated by including dummy variables for each city, technology, and period that compose our panel. As extensively discussed by Wooldridge (2003) and Cameron *et al.* (2011), standard errors should be adjusted for clustering when the errors are correlated within groups of observations, such as cities and technologies in our case. Therefore, all the regression results presented in this article are clustered at the city and technology levels[13].

[12] There is an ongoing debate on whether one should use linear probability model (OLS) or nonlinear specifications (logit, probit). Our preferred specification is based on the linear probability model because it has been shown that the parameter estimates of nonlinear models might not be consistent when there are too many "0" in the dependent variable (King and Zeng, 2001), which is the case in this article. Furthermore, we want to control of unobserved time-invariant heterogeneity using a three-way fixed-effects model. Unfortunately, the statistical properties of the coefficients estimated from fixed-effects logit models are still largely unknown in a panel with a small number of periods (Papke and Wooldridge, 2008). The results, however, are robust to econometric specifications using generalized linear models. Logit and probit specifications can be found in Table 4.

[13] To compute the adjusted standard errors, we use multiway clustering regression techniques implemented in Stata with the "cgmreg" code for the linear probability model, and the "cgmlogit" command for logistic regression (Cameron *et al.*, 2011).

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

Table 2 Summary statistics

Variables	Observed	Mean	Standard deviation	Min	Max
Entry	748,458	0.092585	.2898502	0	1
Exit	213,390	0.3654717	.4815633	0	1
Relatedness density [Hidalgo *et al.*]	961,848	21.58084	29.81881	0	100
Relatedness density [Co-occurrence]	961,848	32.84199	28.20259	0	100
Relatedness density [Hall *et al.*]	961,848	29.96201	28.47284	0	100
Employment [city]	961,848	269,896	648163.6	2630.2	8,538,557
Population density [city]	961,848	2,399,645	2,899,026	3.59798	2790.44
Inventive capacity [city]	961,848	.19754	.21342	0	3.5763
Technological specialization [city]	913,668	15.80558	13.30246	1.30188	63.62291
MSA technological growth rate [city]	904,376	0.0794336	.3963024	−.8743017	.962963
Income per employee [city]	956,592	29.05074	10.86576	10.374	93.57
Nb. Inventors [techno]	961,848	1542.409	2484.688	0	27984
Technological concentration [techno]	913,902	9.756013	11.46718	1.388236	68.73915
Date established [techno]	961,848	1955.249	32.97904	1899	2009
Tech. Class growth rate [techno]	896,700	0.0948035	.4865253	−0.9944946	2.333333

Note: In the econometric estimations presented in this article, employment, income per employee, and Nb, inventors have been log-transformed.

Our panel consists of data for 366 cities (MSAs), 438 technologies (patent technological classes at the three-digit level) over the period of 1976–2010. We average the data[14] over non–overlapping 5-year periods (1976–1980, ..., 2006–2010), denoted by t. To avoid potential endogeneity issues, all the independent variables are lagged by one period[15], so that we have six observations per city–technology pair, resulting in a balanced panel with 961,848 observations[16]. Table 2 provides some summary statistics of the variables used in the econometric analysis[17].

5. Empirical results

In this section, we present the econometric results of the impact of relatedness at the city level on technological change in US cities from 1981 to 2010. We not only

[14] The dummy variables, such as entry, are not averaged.

[15] The first period is only used to construct the independent variables.

[16] In the regression, where entry is the dependent variable, we extracted a subsample from this panel based on the condition that the technology should not belong to the technological portfolio of the city in $t-1$, resulting in 748,458 observations.

[17] The correlation matrix is provided in Table A1.

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

analyze the probability of entry but also the probability of exit of patent technology classes in metropolitan areas.

5.1 Do cities diversify into related technologies?

Table 3 presents the results for the estimation of equation (4). The baseline model (model 1) regresses toward the entry of a given technology in a given city on the density of links around this technology in this city (lagged by one period). Column 1 presents estimation results from pooled OLS[18], while column 5 provides coefficient estimates from the three-way F.E model with all the city and technology variables. In all the different specifications, relatedness density has a positive and significant effect. It indicates that relatedness density has been a crucial driving force behind technological change in US cities for the past 30 years. Relatedness density is not only statistically but also economically significant. If the level of density for a given technology in a given city increases by 10%[19], the probability of the entry of this technology in this city during the next period increases by about 55% (0.051/0.092) in the simplest specification (Table 3; column 1). The economic impact of relatedness density remains stable across the different econometric specifications.

In order to verify that our results are not affected by omitted variables bias, we control for important city and technology characteristics. A second model (model 2), reported in column 2, includes variables that capture the heterogeneity of cities that might influence technological change. As expected, the economic size of cities (*employment*), the ratio of inventors to employees (*inventive capacity*), and the growth rate of the number of inventors (*MSA technological growth rate*) play a positive and significant role on the entry of new technologies. Population density also has a positive impact, but the coefficient is not statistically significant. Our results also confirm the idea that cities characterized by a very specialized technological structure (*technological specialization*) are less prone to technological change. A more counter-intuitive finding, however, is the negative role played by the economic wealth of the city (*income per employee*). It might be explained by the fact that once one included the variables discussed above, income per employee does not reflect the inventive capacity of cities. A third model (model 3), reported in column 3, includes variables that capture the heterogeneity of technological classes that might influence their general entry in cities. It is not surprising that large technological classes, i.e. with

[18] In all pooled OLS specifications, the independent variables are mean-centered in order to facilitate the direct interpretation of the results that are proposed in the text. In these specifications, the constant term then reflects the expected probability of entry/exit when the independent variables are set to their mean (the intercept is equal to the mean of the dependent variable presented in Table 2).

[19] The coefficients provided in the table indicate the impact of a 1-unit change on the probability of entry, and the density variable is expressed in percentage. Therefore, when we write that the density variable increases by 10%, we mean by 10 points, i.e. from a level of 25% to a level of 35%, for instance.

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

Table 3 Emergence of new technologies in US cities (1981–2010)

Dependent variable is: $Entry_t$	Model 1 Rel. density	Model 2 City variables	Model 3 Tech. variables	Model 4 Full model	Model 5 Full model (F.E.)
Relatedness $density_{t-1}$	0.00515979**			0.00373407**	0.00271463**
	(0.00012770)			(0.00014135)	(0.00016884)
Log $(Employment)_{t-1}$		0.04934166**		0.03611889**	0.04633250**
		(0.00286818)		(0.00247147)	(0.00782869)
Population $density_{t-1}$		0.00001106		0.00002520**	−0.00021341**
		(0.00000997)		(0.00000843)	(0.00003836)
Inventive $capacity_{t-1}$		0.07718815**		0.03883926**	−0.08487966**
		(0.01294204)		(0.0078352020)	(0.01505564)
Tech. $Specialization_{t-1}$		−0.00089296**		−0.00047160**	0.00005120
		(0.00011548)		(0.00009315)	(0.00011022)
MSA growth $rate_{t-1}$		0.04443962**		0.04032813**	0.00865397**
		(0.00355534)		(0.00353667)	(0.00298386)
Log (Income per $employee)_{t-1}$		−0.07584685**		−0.10127439**	0.00368879
		(0.00441610)		(0.00538561)	(0.01663469)
Log (Nb. $Inventors)_{t-1}$			0.02658895**	0.02324554**	0.00159990
			(0.00197752)	(0.00183672)	(0.00246612)
Tech. $concentration_{t-1}$			−0.00102840**	−0.00010693	0.00041990 *
			(0.00014936)	(0.00011541)	(0.00016760)
Date $established_{t-1}$			−0.00056684**	−0.00042520**	−0.00330620**
			(0.00007012)	(0.00005456)	(0.00017699)
Tech. growth $rate_{t-1}$			0.01423964**	0.02183910**	0.01141729**
			(0.00233334)	(0.00285492)	(0.00260757)
Constant	0.09258502**	0.09296771**	0.09019069**	0.08909252**	0.11108572**
	(0.00194271)	(0.00378306)	(0.00398429)	(0.00183778)	(0.01040890)
City F.E.	No	No	No	No	Yes
Technology F.E.	No	No	No	No	Yes
Period F.E.	No	No	No	No	Yes
R^2	0.11	0.04	0.02	0.13	0.16
N	748,458	653,660	656,618	572,550	572,550

Notes: The dependent variable entry = 1 if a given technology ($n = 438$) enters in the technological portfolio of a given city ($n = 366$) during the corresponding 5-year window ($n = 6$), and 0 otherwise. The independent variables are centered around their mean. Coefficients are statistically significant at the $*P < 0.05$ and $**P < 0.01$ level. Heteroskedasticity-robust standard errors (clustered at the city and technology level) are in parentheses.

a large pool of inventors are more likely to enter in any US metropolitan area (*Nb. Inventors*), especially if the production of knowledge in this technological class is growing (*Tech. Class growth rate*). On the contrary, older technologies (*date established*) and technologies that are very much concentrated in space are significantly less likely to be developed by many different cities in the future. Overall, almost all

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

the variables explain an important part of the variation in terms of entry of new technologies, and therefore they are important predictors of technological change.

In the full model specification (model 4), we tested whether the effect of relatedness density was affected by these important features of technologies and cities. Column 4 presents estimation results from a pooled OLS, while column 5 presents the complete estimations results from equation (4), i.e. including relatedness density; city and technology time-varying variables; and F.E for cities, technologies, and time. Interestingly, the coefficient for density remains highly significant, but its magnitude slightly decreases with the addition of these control variables. When F.E are included, the rate of entry increases by approximately 30% for a 10% increase in the level of density in city–technology pairs[20] (0.027/0.089).

5.2 Robustness analysis

In Table 4, we present alternative econometric specifications to test the robustness of the relationship of interest, i.e. the effect of relatedness density on technological change in cities. We run three different sets of robustness checks: (i) using alternative measures of technological relatedness as independent variables, (ii) excluding observations with extreme values (i.e. outliers), and (iii) using alternative econometric methods to the linear probability model. The results reported[21] in Table 4 show that the positive and significant impact of relatedness density on the probability of entry is robust to these alternative specifications.

First, we verify that our results are not driven by the technological relatedness measure we used (Table 4; columns 1–2). Therefore, we estimated equation (4) (see specification in Table 3; column 5) by using the Hall *et al.* (2001) patent classification and a normalized co-occurrence analysis. The coefficient[22] on relatedness density is smaller in those alternative specifications, but remains statistically and economically significant. When density increases by 10%, the probability of entry increases by approximately 15% using the Hall *et al.* (2001) classification and by 20% using the co-occurrences analysis.

[20] One of the referees interestingly pointed out that the impact of density might not be equally important for small, medium-sized, and/or large cities. To explore this question, we divided the cities in our sample into three equal groups based on their number of employees. Computing standardized coefficients, we found that the impact of density is equally important for medium and large cities, but slightly lower for small cities. It suggests that relatedness at the city level still requires some critical mass to lead to technological diversification.

[21] Robustness checks have been carried out for all models presented in Table 3, and our main variable of interest remains positive and significant. Since the fixed-effects model is probably the most conservative, we focus on this specification to present the robustness checks in Table 4.

[22] We computed the standardized coefficient for the three different density measures, and the coefficients are smaller in magnitude when using the Hall *et al.* (2001) classification and co-occurrence analysis than when using the Hidalgo *et al.* (2007) method.

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

Table 4 Entry of new technologies in U.S. cities - Robustness check

Dependent variable is: $Entry_t$	Alternative relatedness measures		Outliers analysis			GLM specifications	
	Model COOC [F.E.]	Model USPTO [F.E.]	w/o top density[a]	w/o top cities[b]	w/o top techno[c]	Logistic regression	Probit regression
$Density_{t-1}$ [baseline]			0.00224635** (0.00016733)	0.00264742** (0.00018286)	0.00239119** (0.00017202)	0.0216442** (0.0002433)	0.0125646** (0.0001334)
$Density_{t-1}$ [COOC]	0.00184525** (0.00016940)						
$Density_{t-1}$ [USPTO]		0.00142651** (0.00014815)					
City controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tech. controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City F.E.	Yes	Yes	Yes	Yes	Yes	No	No
Technology F.E.	Yes	Yes	Yes	Yes	Yes	No	No
Period F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2/Pseudo R^2	0.15	0.15	0.11	0.15	0.14	0.19	0.19
N	572,550	572,550	495,077	515,350	514,091	572,550	572,550

Notes: The dependent variable entry = 1 if a given technology ($n = 438$) enters in the technological portfolio of a given city ($n = 366$) during the corresponding 5-year window ($n = 6$), and 0 otherwise. Coefficients are statistically significant at the $*P < 0.05$ and $**P < 0.01$ level. Heteroskedasticity-robust standard errors (clustered at the city–technology level for the logistic and probit regression; clustered at the city and technology level in all other regressions) are in parentheses.

[a]The top 10% of the city–technology pairs with the highest density are dropped, [b]The top 10% of the cities that experienced the highest number of technology entry are dropped, [c]The top 10% of the technologies that entered cities the most frequently are dropped.

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

Second, we check that our results were not driven by extreme values at the top decile level (Table 4; columns 3–5). Using our baseline measure of technological relatedness (Hidalgo *et al.*, 2007), we estimated equation (4) (Table 3; column 5) by removing the top 10% of the city–technology pairs with the highest density (Table 4; column 3), by removing the top 10% of the cities that experienced the highest number of technology entry (Table 4; column 4), and finally by removing the top 10% of the technologies that entered cities the most frequently (Table 4; column 5). None of these alternative specifications with altered data samples seem to substantially affect the statistical or economic effect of relatedness density.

Third, we estimated equation (4) with generalized linear models (GLM), i.e. logit and probit (Table 4; columns 6 and 7). In the article, we focused on linear probability models, but since the dependent variable (entry) is dichotomous, we also check that our results are robust to traditional GLM specifications. The last two columns in Table 4 show results from logistic and probit regressions, and they confirm the positive and significant impact of density of related technologies on the probability of entry.

The alternative measures of technological relatedness, and also estimates from data samples without extreme values and alternative econometric models, all support our key findings. On top of that, we also estimated models by using density from weighted relatedness matrices, by using all classes listed on patent documents to construct the relatedness space and the corresponding entry variables, by using inventor shares to localize patents instead of the primary inventor, and by constraining the entry events to technologies in which cities have a comparative advantage[23]. These additional analyses do not affect the results presented here and suggest that our statistically and economically significant positive relationship between relatedness density and the probability of entry is robust to several key econometric specifications.

5.3 Does relatedness density prevent the exit of technologies?

But technological change is not only about entry of new technologies within cities. In fact, technological change can be understood as a process of creative destruction, in which the exit of existing technologies also contributes to change the technological landscape of cities. Table 5 reports estimation results where the dependent variable "exit" is used instead of the dependent variable "entry." The results indicate that relatedness density has a negative and significant impact on the exit of technologies. If the level of density for a given technology in a given city increases by 10%, the probability of exit of this technology in this city during the next period decreases by about 8%–17%, depending on the econometric specifications. The results concerning

[23] We would like to thank one of the editors for suggesting these additional robustness checks. The results of these analyses are not reported in the article but available upon request from the authors.

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

Table 5 Exit of technologies in US cities (1981–2010)

Dependent variable is: $Exit_t$	Model 1 Rel. density	Model 2 City variables	Model 3 Tech. variables	Model 4 Full model	Model 5 Full model (F.E.)
Relatedness $density_{t-1}$	−0.00646272**			−0.00384300**	−0.00287999**
	(0.00013398)			(0.00022311)	(0.00021200)
Log $(Employment)_{t-1}$		−0.10857437**		−0.06943327**	−0.08359852**
		(0.00614202)		(0.00626204)	(0.01651044)
Population $density_{t-1}$		−0.00003837*		−0.00006553**	−0.00011335*
		(0.00001950)		(0.00001364)	(0.00004718)
Inventive $capacity_{t-1}$		−0.16931248**		−0.11188970**	−0.02739567**
		(.05336078)		(.02941733)	(.00841076)
Tech. $Specialization_{t-1}$		0.00437970**		0.00180088**	−0.00056492
		(0.00061919)		(0.00040634)	(0.00042826)
MSA growth $rate_{t-1}$		−0.16187457**		−0.15036339 **	−0.01352593
		(0.00828661)		(0.00882790)	(0.00966971)
Log (Income per $employee)_{t-1}$		0.22689471**		0.31767021**	0.04962913
		(0.01306891)		(0.01236049)	(0.03082188)
Log (Nb. $Inventors)_{t-1}$			−0.04660531**	−0.09098814**	−0.05541312**
			(0.00593058)	(0.00406299)	(0.00624571)
Tech. $concentration_{t-1}$			0.00418752**	−0.00137922**	−0.00200006**
			(0.00047497)	(0.00043938)	(0.00058743)
Date $established_{t-1}$			−0.00018739	0.00022470*	0.00233776**
			(0.00011545)	(0.00010297)	(0.00022809)
Tech. growth $rate_{t-1}$			−0.06741134**	−0.05102216**	−0.01451667 *
			(0.00590281)	(0.00701074)	(0.00652948)
Constant	0.36547167**	0.36534949**	0.36590647**	0.36470402**	0.54798934**
	(0.00609779)	(0.00841466)	(0.01460965)	(0.00426248)	(0.03747886)
City F.E.	No	No	No	No	Yes
Technology F.E.	No	No	No	No	Yes
Period F.E.	No	No	No	No	Yes
R^2	0.19	0.15	0.03	0.25	0.30
N	213,390	202,584	201,286	191,313	191,313

Notes: The dependent variable entry = 1 if a given technology ($n = 438$) exits the technological portfolio of a given city (n = 366) during the corresponding 5-year window ($n = 6$), and 0 otherwise. The independent variables are centered around their mean. Coefficients are statistically significant at the $^*P < 0.05$ and $^{**}P < 0.01$ level. Heteroskedasticity-robust standard errors (clustered at the city and technology level) are in parentheses.

city and technology characteristics are also consistent with our expectations. Cities with a high economic potential are more likely to prevent the exit of technologies, while economically important technologies are less likely to exit in all the cities. What should be noticed, however, is that the relative economic importance of relatedness

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

density compared with city and technology characteristics seems to be smaller to explain variations in the exit than the entry of technologies.

6. Discussion and Concluding Remarks

In this article, we found evidence that the rise and fall of technological knowledge, as proxied by the entry and exit of patent technology classes in cities, is conditioned by the existing technological knowledge base of cities. Analyzing the long-term evolution of patent portfolios of 366 US cities during the period of 1976–2010, we found that a new technology is more likely to enter a city when technologically related to other technologies in that city, and an existing technology had a higher probability to exit a city when it was not, or poorly, related to other technologies in that city. These results indicate that technological relatedness was a driving force behind technological change in US cities in the past 30 years, and that the long-term evolution of the technological urban landscape is subject to path dependency.

These findings call for further research. As new technologies emerge systematically in cities with related technologies, this suggests that new technologies are all recombinations of existing technologies. While this might be true for a large fraction of new technologies, it is not necessarily true for all of them. In fact, some new technologies (patents) are truly novel, with few or no related technologies on which these are built during their time of emergence (Dahlin and Behrens, 2005; Castaldi and Los, 2007; Krafft *et al.*, 2011). From a geographical perspective, it would be interesting to investigate where radical technologies and new technological trajectories come into being. For instance, do these need highly diversified cities, instead of technologically specialized cities (Duranton and Puga, 2000)?

Another issue to be taken up in future research is whether new patents actually build on related patents in cities. By looking at the set of (related) patents at the city level, we did not investigate the extent to which a new patent that is new for a city actually cites other patents in related technology classes from the same city. This would provide evidence at the level of patents (rather than at the level of cities) that invention activity actually builds on related knowledge at the city level. This would also shed light on the importance of knowledge flows from other cities, as patents might draw on and cite related patents from other cities. Another issue is the selection of the relatedness indicator to study urban diversification. We made use of co-occurrence analysis based on the frequency of combinations of patent classes within the same cities. Other scholars like Leten *et al.* (2007), Van der Wouden (2012), and Rigby (2013) have used alternative indicators to measure relatedness between patent classes, such as patent citations. However, recent studies, such as those by Rigby (2013), have shown that findings are unlikely to change when using such alternative measures of technological relatedness. As a robustness check, we made use of two alternative measures of relatedness (i.e. normalized co-occurrences and patent classes

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

belonging to the same two-digit category), and our findings with respect to relatedness density remained statistically significant. Other studies on regional diversification have used other measures of relatedness based on the intensity of input–output linkages between industries (Essletzbichler, 2013) or on co-occurrence analysis of product categories either within plants (Neffke *et al.*, 2011) or within countries (Boschma *et al.*, 2013). Although these studies are very different in terms of their relatedness measure, the use of spatial units and methodologies, the period covered, and the selection/measurement of the dependent and independent variables, they also found evidence of relatedness at the regional scale driving regional diversification.

A final issue deserves attention in future research. In this article, we explored the extent to which the entry of a new technology depends on other technologies to which it is related at the city level. However, we did not explore other dimensions that might be crucial in the process of technological diversification, such as institutional preconditions at the city level (Strambach, 2010). In fact, such a study would shed light on the extent to which related technologies draw on and require similar sets of urban institutions, which could provide an additional explanation for the fact that related technologies tend to benefit from each other's copresence in cities.

Acknowledgements

This work is part of a project granted under the Open Research Area in Europe for the Social Sciences (ORA): '*Territories and technologies in an unstable knowledge economy: An evolutionary framework of regional resilience*'. We would like to thank the Nederlandse Organisatie voor Wetenschappelijk Onderzoek, (NWO), Agence Nationale de la Recherche (ANR), Economic and Social Research Council (ESRC) and Deutsche Forschungsgemeinschaft (DFG) for funding this project.

References

Alcácer, J. and M. Gittelman (2006), 'Patent citations as a measure of knowledge flows: the influence of examiner citations,' *The Review of Economics and Statistics*, **88**, 774–779.

Antonelli, C. (1995), *The Economics of Localized Technological Change and Industrial Dynamics*. Dordrecht: Kluwer.

Archibugi, D. and M. Pianta (1992), 'Specialization and size of technological activities in industrial countries: the analysis of patent data,' *Research Policy*, **21**, 79–93.

Atkinson, A. and J. Stiglitz (1969), 'A new view of technical change,' *Economic Journal*, **79**, 573–578.

Autant-Bernard, C. (2001), 'The geography of knowledge spillovers and technological proximity,' *Economics of Innovation and New Technology*, **10**, 237–254.

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

Bathelt, H., M. P. Feldman and D. F. Kogler (eds) (2011), *Beyond Territory. Dynamic Geographies of Knowledge Creation, Diffusion and Innovation*. Routledge: London/New York.

Beaudry, C. and A. Schiffauerova (2009), 'Who's right, Marshall or Jacobs? The localization versus urbanization debate,' *Research Policy*, **38**, 318–337.

Bettencourt, L., J. Lobo and D. Strumky (2007), 'Invention in the city: Increasing returns to patenting as a scaling function of metropolitan size,' *Research Policy*, **36**, 107–120.

Boschma, R. and K. Frenken (2011), 'Technological relatedness and regional branching,' in H. Bathelt, M.P. Feldman and D.F. Kogler (eds), *Beyond Territory. Dynamic Geographies of Knowledge Creation, Diffusion and Innovation*. Routledge: London and New York, pp. 64–81.

Boschma, R., A. Minondo and M. Navarro (2013), 'The emergence of new industries at the regional level in Spain. A proximity approach based on product-relatedness,' *Economic Geography*, **89**(1), 29–51.

Boschma, R. A. and R. Wenting (2007), 'The spatial evolution of the British automobile industry. Does location matter?' *Industrial and Corporate Change*, **16**, 213–238.

Breschi, S., S. Lissoni and F. Malerba (2003), 'Knowledge-relatedness in firm technological diversification,' *Research Policy*, **32**, 69–87.

Buenstorf, G. and C. Guenther (2011), 'No place like home? Relocation, capabilities, and firm survival in the German machine tool industry after World War II,' *Industrial and Corporate Change*, **20**, 1–28.

Buenstorf, G. and S. Klepper (2009), 'Heritage and agglomeration: the Akron tyre cluster revisited,' *The Economic Journal*, **119**, 705–733.

Cameron, A. C., J. B. Gelbach and D. L. Miller (2011), 'Robust inference with multiway clustering,' *Journal of Business and Economic Statistics*, **29**, 238–249.

Cantwell, J. and G. Vertova (2004), 'Historical evolution of technological diversification,' *Research Policy*, **33**, 511–529.

Castaldi, C. and B. Los (2007), 'International technological specialization in important innovations: some industry-level explorations,' *Working paper*. University of Groningen: Groningen.

Chapman, K. (1991), *The International Petrochemical Industry. Evolution and Location*. Basil Blackwell: Oxford.

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

Dahlin, K. B. and D. M. Behrens (2005), 'When is an invention really radical? Defining and measuring technological radicalness,' *Research Policy*, **34**, 717–737.

Dosi, G. (1982), 'Technological paradigms and technological trajectories. a suggested interpretation of the determinants and directions of technical change,' *Research Policy*, **29**, 497–529.

Duranton, G. and D. Puga (2000), 'Diversity and specialisation in cities: why, where and when does it matter?' *Urban Studies*, **37**, 533–555.

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

Ejermo, O. (2005), 'Technological diversity and Jacobs' externality hypothesis revisited,' *Growth and Change*, **36**, 167–195.

Eriksson, R. H. (2011), 'Localized spillovers and knowledge flows–How does proximity influence the performance of plants,' *Economic Geography*, **87**, 127–152.

Essletzbichler, J. (2013), 'Relatedness, industrial branching and technological cohesion in US metropolitan areas,' *Regional Studies*, forthcoming, DOI: 10.1080/00343404.2013.806793.

Feldman, M. P. and D. B. Audretsch (1999), 'Innovation in cities: science-based diversity, specialization, and localized competition,' *European Economic Review*, **43**, 409–429.

Fleming, L. and O. Sorenson (2001), 'Technology as a complex adaptive system: evidence from patent data,' *Research Policy*, **30**, 117–132.

Frenken, K. and R. Boschma (2007), 'A theoretical framework for evolutionary economic geography: industrial dynamics and urban growth as a branching process,' *Journal of Economic Geography*, **7**, 635–649.

Frenken, K., F. G. Van Oort and T. Verburg (2007), 'Related variety, unrelated variety and regional economic growth,' *Regional Studies*, **41**, 685–697.

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.

Glaeser, E. L. (2005), 'Reinventing Boston: 1630-2003,' *Journal of Economic Geography*, **5**, 119–153.

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

Hall, B., A. Jaffe and M. Trajtenberg (2001), 'The NBER Patent citation data file: lessons, insights, and methodological tools,' *National Bureau of Economic Research Working Paper 8498*. MIT Press: Cambridge, MA.

Hall, P. (1998), *Cities in Civilization*. Pantheon Books: New York.

Hassink, R. (2010), 'Locked in decline? On the role of regional lock-ins in old industrial areas,' in R.A. Boschma and R. Martin (eds), *Handbook of Evolutionary Economic Geography*. Edward Elgar: Cheltenham.

Hausmann, R. and C. A. Hidalgo (2010), 'Coun,' try diversification, product ubiquity, and economic divergence,' *CID Working Paper No. 201*. Center for International Development, Harvard University.

Hausmann, R. and B. Klinger (2007), 'The Structure of the Product Space and the Evolution of Comparative Advantage,' *CID Working Paper 146*. Center for International Development, Harvard University.

Heiner, R. A. (1983), 'The origin of predictable behaviour,' *The American Economic Review*, **73**, 560–595.

Hidalgo, C. A., B. Klinger, A. -L. Barabasi and R. Hausmann (2007), 'The product space conditions the development of nations,' *Science*, **317**, 482–487.

iCancho, R. F. and R. V. Solé (2001), 'The small world of human language,' *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **268**, 2261–2265.

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

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

King, G. and L. Zeng (2001), 'Logistic regression in rare events data,' *Political Analysis*, **9**, 137–163.

Klepper, S. (2007), 'Disagreements, spinoffs, and the evolution of Detroit as the capital of the U.S. automobile industry,' *Management Science*, **53**, 616–631.

Kogler, D. F., D. L. Rigby and I. Tucker (2013), 'Mapping knowledge space and technological relatedness in U.S. cities,' *European Planning Studies*, **21**(9), 1374–1391.

Krafft, J., F. Quatraro and P. Saviotti (2011), 'The knowledge-base evolution in biotechnology: a social network analysis,' *Economics of Innovation and New Technology*, **20**, 445–475.

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

Lawson, C. (1999), 'Towards a competence theory of the region,' *Cambridge Journal of Economics*, **23**, 151–166.

Leten, B., R. Belderbos and B. Van Looy (2007), 'Technological diversification, coherence, and performance of firms,' *Journal of Product Innovation and Management*, **24**, 567–579.

Lobo, J. and D. Strumsky (2008), 'Metropolitan patenting, inventor agglomeration and social networks: a tale of two effects,' *Journal of Urban Economics*, **63**, 871–884.

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

Malmberg, A. and P. Maskell (1997), 'Towards an explanation of regional specialization and industry agglomeration,' *European Planning Studies*, **5**, 25–41.

Maskell, P. and A. Malmberg (1999), 'The competitiveness of firms and regions. Ubiquitification and the importance of localized learning,' *European Urban and Regional Studies*, **6**, 9–25.

Maskell, P. and A. Malmberg (2007), 'Myopia, knowledge development and cluster evolution,' *Journal of Economic Geography*, **7**, 603–618.

Neffke, F. (2009), 'Productive places the influence of technological change and relatedness on agglomeration externalities,' PhD thesis, Utrecht University: Utrecht.

Neffke, F., M. Henning and R. Boschma (2011), 'How do regions diversify over time? Industry relatedness and the development of new growth paths in regions,' *Economic Geography*, **87**, 237–265.

Nelson, R. R. and S.G. Winter (1982), *An Evolutionary Theory of Economic Change*. The Belknap Press: Cambridge, MA and London.

Nesta, L. (2008), 'Knowledge and productivity in the world's largest manufacturing operations,' *Journal of Economic Behavior and Organization*, **67**, 886–902.

O'hUallachain, B. and D. Lee (2010), 'Technological variety and specialization in urban invention,' *Regional Studies*, **45**, 67–88.

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

O'hUallachain, B. and T. F. Leslie (2007), 'Rethinking the regional knowledge production function,' *Journal of Economic Geography*, **7**, 737–752.

Office of Management and Budget (2010), Standards for Delineating Metropolitan and Micropolitan Statistical Areas. 75 Federal Register 37246, Monday, June 28, 2010.

Paci, R. and S. Usai (1999), 'Externalities, knowledge spillovers and the spatial distribution of innovation,' *GeoJournal*, **49**, 381–390.

Papke, L. E. and J. M. Wooldridge (2008), 'Panel data methods for fractional response variables with an application to test pass rates,' *Journal of Econometrics*, **145**, 121–133.

Pavitt, K. (1985), 'Patent statistics as indicators of innovative activities: pPossibilities and problems,' *Scientometrics*, **7**, 77–99.

Penrose, E. (1959), *The Theory of the Growth of the Firm*. Oxford University Press: Oxford.

Piscitello, L. (2004), 'Corporate diversification, coherence and economic performance,' *Industrial and Corporate Change*, **13**, 757–787.

Quatraro, F. (2010), 'Knowledge coherence, variety and economic growth: manufacturing evidence from Italian regions,' *Research Policy*, **39**, 1289–1302.

Rigby, D. L. (2013), 'Technological relatedness and knowledge space: entry and exit of US cities from patent classes,' *Regional Studies*, forthcoming, doi: 10.1080/00343404.2013.854878.

Rigby, D. L. and J. Essletzbichler (1997), 'Evolution, process variety, and regional trajectories of technological change in US manufacturing,' *Economic Geography*, **73**, 269–284.

Scherer, F. M. (1984), *Innovation and Growth: Schumpeterian Perspectives*. MIT Press: Cambridge, MA.

Sonn, J. and M. Storper (2008), 'The increasing importance of geographical proximity in knowledge production: an analysis of U.S. patent citations,' *Environment and Planning A*, **40**, 1020–1039.

Storper, M. (1995), 'The resurgence of regional economies, ten years later: the region as a nexus of untraded interdependencies,' *European Urban and Regional Studies*, **2**, 191–221.

Strambach, S. (2010), 'Path dependency and path plasticity. The co-evolution of institutions and innovation - the German customized business software industry,' in R. A. Boschma and R. Martin (eds), *Handbook of Evolutionary Economic Geography*. Edward Elgar: Cheltenham.

Strumsky, D., J. Lobo and S. Van der Leeuw (2012), 'Using patent technology codes to study technological change,' *Economics of Innovation and New Technology*, **21**, 267–286.

Teece, D., R. Rumelt, G. Dosi and S. Winter (1994), 'Understanding corporate coherence: theory and evidence,' *Journal of Economic Behavior and Organization*, **23**, 1–30.

Van der Wouden, F. (2012), 'The role of relatedness in economic variety on knowledge production of U.S. cities between 1975 and 2005,' MA thesis, Department of Economic Geography, Utrecht University.

Wooldridge, J. M. (2003), 'Cluster-sample methods in applied econometrics,' *American Economic Review*, **93**, 133–138.

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015

Appendix

Table A1 Correlation matrix

	Entry	Exit	Rel. D1	Rel. D2	Rel. D3	Emp.	Pop. Dens	Inv. cap.	Spe.	TGR	Income	Nb. inv	Conc.	Date	TGR
Entry	1.0000														
Exit	.	1.0000													
Relatedness density [Hidalgo *et al.*]	0.3346	−0.4395	1.0000												
Relatedness density [Co-occurrence]	0.1933	−0.3922	0.6206	1.0000											
Relatedness density [Hall *et al.*]	0.2049	−0.3686	0.6311	0.8212	1.0000										
Employment [city]	0.1751	−0.3441	0.5978	0.7820	0.7388	1.0000									
Population density [city]	0.1052	−0.2301	0.4417	0.5504	0.5265	0.6457	1.0000								
Inventive capacity [city]	0.0848	−0.1147	0.3414	0.4604	0.4306	0.2587	0.2656	1.0000							
Technological specialization [city]	−0.1394	0.2807	−0.4162	−0.5813	−0.5421	−0.5682	−0.3619	−0.3565	1.0000						
MSA technological growth rate [city]	0.0522	−0.0755	0.1115	0.1627	0.1484	0.1415	0.0564	0.2608	−0.1329	1.0000					
Income per capita [city]	−0.0170	0.0060	0.1514	0.2067	0.1861	0.2348	0.2028	0.3401	−0.0952	0.4873	1.0000				
Nb. inventors [techno]	0.1436	−0.1652	0.3454	0.0051	0.0506	0.0219	0.0107	0.0404	−0.0034	0.0617	0.1396	1.0000			
Technological concentration [techno]	−0.1115	0.1275	−0.2715	0.0440	−0.0333	0.0060	0.0030	0.0096	0.0001	0.0232	0.0430	−0.5216	1.0000		
Date established [techno]	−0.0124	−0.0524	−0.0150	−0.0456	0.0004	−0.0000	−0.0000	−0.0000	−0.0000	−0.0010	−0.0000	0.2026	−0.1216	1.0000	
Tech. Class growth rate [techno]	0.0622	−0.1055	0.0969	0.0050	0.0417	0.0354	0.0141	0.0428	−0.0118	0.3242	0.2561	0.3244	−0.1613	0.1529	1.0000

Note: To correspond with the variables used in the econometric estimations presented in the article, employment, income per capita, and Nb, inventors have been log-transformed before computation of the correlation matrix (Pearson's correlation coefficient).

Downloaded from http://icc.oxfordjournals.org/ at Universiteitsbibliotheek Utrecht on November 24, 2015