Knowledge specialization, knowledge brokerage, and the uneven growth of technology domains

Forthcoming in SOCIAL FORCES
(Note: This version has not been copyedited by Social Forces)

Gianluca Carnabuci[a]
University of Lugano

Jeroen Bruggeman
University of Amsterdam

Acknowledgements: The authors are grateful to Bernhard Kittel, Rob Mokken, Willem Saris, Joel Podolny, Frans Stokman, Bart Verspagen, the editor, and three anonymous reviewers, for their useful comments and suggestions. This research was supported by the Netherlands Organization for Scientific Research, NWO grant #401-01-098. The first author is grateful to the Research Division of SDA Bocconi School of Management for its support.

[a] **Corresponding author**. Institute of Management and Organization, University of Lugano, Via Buffi 13, 6903, Lugano, Switzerland. Email: gianluca.carnabuci@lu.unisi.ch

Abstract: Why do certain domains of knowledge grow fast while others grow slowly or stagnate? Two distinct theoretical arguments hold that knowledge growth is enhanced by, respectively, knowledge specialization and knowledge brokerage. Based on the notion of recombinant knowledge growth, we show that specialization and brokerage are opposing modes of knowledge generation, the difference between them lying in the extent to which homogeneous versus heterogeneous input ideas get creatively recombined. Accordingly, we investigate how both modes of knowledge generation can enhance the growth of technology domains. To address this question, we develop an argument that reconciles both specialization and brokerage into a dynamic explanation. Our contention is that specializing in an increasingly homogeneous set of input ideas is both more efficient and less risky than brokering knowledge. Nevertheless, specializing implies progressively exhausting available recombinant possibilities, while brokerage creates new ones. Hence, technology domains tend to grow faster when they specialize, but the more specialized they become, the more they need knowledge brokerage to grow. We cast out our argument into five hypotheses that predict how growth rates vary across technology domains. Based on all technological knowledge patented in the USA from 1975 to 1999, our hypotheses are corroborated.

Introduction

The ability of human societies to transform economic inputs into outputs of greater value, and hence to generate material wealth, rests on their technological knowledge (Jones 2005). In contrast to preliterate societies, where information about inventions circulated mostly locally by word of mouth (Diamond 2001), a distinguishing trait of the knowledge-based economy[1] is that a considerable share of the existing technological knowledge is a *public good* (Mokyr 2002). Technological knowledge is a public good to the extent that it is clearly explicated and codified, and it is made widely available through media accessible at negligible costs, such as patents and technical literature (Arrow 1962). As most scholars agree, the growth rate of public technological knowledge has increased progressively during the last couple of centuries, yielding an unprecedented yet unrelenting pace of economic growth (Jones 2005). The mechanisms and dynamics driving the growth of public technological knowledge, however, are still poorly understood.

To begin covering this gap, in this paper we investigate why some domains of technology grow fast while others grow slowly or stagnate (Nelson 2003). We understand the growth of a technology domain as the extent to which new inventions are generated within that domain that engender useful new ideas and applications. Hence, our definition is in line with the notion that the growth of knowledge reflects both the amount of new ideas generated *and* the extent to which these ideas have gained acceptance and public recognition (Simonton 2000; Walberg 1998; Weisberg 1993). To explain growth differentials across technology domains, we analyze the intertwined effects of two modes of knowledge generation that are commonly assumed to enhance knowledge growth in general. First, it has been argued that knowledge *specialization* facilitates the progress of knowledge by increasing the efficiency of the knowledge generation process, a notion that is at the core both of economic theory (e.g. Smith, 1776, Walker 1867, Young 1928, Marshall 1936) and of bounded-rationality theories of learning (Simon 1977). Second, the argument that *brokerage* of knowledge across disparate domains yields novelty and thereby boosts knowledge generation, has recently become widespread in sociological publications (e.g. Burt 2004, Stuart and Podolny 1996, Sutton and Hargadon 1996).

Over the years, both the specialization and the brokerage arguments have gained substantial empirical support and many followers. The theoretical relation between knowledge specialization and knowledge brokerage has not yet been explicated in any detail, though, and this task involves conceptual challenges that we hope our paper will help to overcome. To relate the specialization and brokerage arguments, we build on the concept of "recombinant knowledge growth" (Weitzman 1996, 1998) by taking the position that, whether generated through knowledge specialization or through knowledge brokerage, new knowledge always derives from recombinations of existing knowledge (Schumpeter 1939, Usher 1954, Weitzman 1998, Fleming 2001, Nolan and Lenski 2006). While the mechanism of knowledge recombination is common to both knowledge specialization and knowledge brokerage, we contend that the difference between these two modes of knowledge generation lies in the extent to which creative recombinations build on homogeneous rather than heterogeneous input ideas. Hence, a technology domain is specialized to the extent that its new knowledge builds on a homogeneous set of closely related ideas; by the same token, the more that new knowledge builds on a wide spectrum of heterogeneous knowledge, the more that domain is brokering.

Seen from the perspective of recombinant knowledge growth, knowledge specialization and knowledge brokerage are therefore endpoints of one conceptual continuum, rather than distinct concepts. That is, the higher the degree of specialization of a technology domain, the lower its degree of brokerage, and vice-versa. But if a specialized domain is the opposite of a brokering domain, it follows that the growth of technological knowledge cannot be faster in both specialized and brokering domains at the same time, as the specialization and brokerage arguments would suggest if taken individually. To resolve this apparent paradox, we will develop a theory that integrates the specialization and brokerage arguments into a dynamic explanation. The core of our thesis is that knowledge brokerage generates new opportunities for knowledge recombination, while these opportunities are more efficiently and more securely exploited through increasing knowledge specialization. As a result, we will argue and show that when technology domains undergo a process of increasing specialization they tend to grow faster than when they fail to specialize (or even they de-specialize); however, the higher is the level of specialization a technology domain has reached, the more its future progress requires brokering knowledge from yet unrelated domains in a de-specializing, path-breaking, fashion. In addition, we will show that the degree of specialization reached by a domain impinges on how volatile its future growth rate is going to be, specialized domains being far more predictable than brokering ones.

To further explain and to empirically demonstrate our arguments, we proceed as follows. In Section 1, we outline the view of recombinant knowledge growth, arguing how we apply it to the aggregate level of technology domains. Subsequently, in Section 2, we explicate recombinant knowledge growth as a network model; we then define domains' specialization and brokerage on the basis of that model. Against the backdrop of our network model, in Section 3 we develop our theory and derive five hypotheses concerning their effects on the growth of technology domains. To test our hypotheses, we look at all knowledge patented in the USA over a quarter of a century. We employ the technological classifications of the United States Patent and Trademark Office (henceforth, USPTO) as proxies for technology domains. Furthermore, we use backward patent citations to map the dynamics of the specialization and brokerage of technology domains, and we use changes in the number of forward citations over time to gauge domains' growth. Both our empirical data and the operationalization of our variables are elaborated in Section 4. In Section 5, we report the statistical methods and the results of our analyses. Finally, in Section 6, we discuss the main contributions, implications, and limits of our study.

1. The salience of technology domains in recombinant knowledge growth

Public knowledge tends to diffuse in spite of geographic, organizational and social barriers, and it does not deplete with usage (Arrow 1962). Furthermore, new ideas spring from the novel combination of earlier ideas, and thus the input for future knowledge is existing knowledge (Gilfillan 1935; Schumpeter 1939; Usher 1954; Romer 1993).[2] In an influential pair of papers inspired by these considerations, economist Martin Weitzman proposed to conceive the growth of public knowledge as a process of "recombinant knowledge growth", which he represented mathematically as a combinatorial process (Weitzman 1996, 1998). Weitzman's modelling strategy captures elegantly what is arguably the most distinctive trait of the knowledge generation process, namely that because each new idea can potentially be recombined to produce multiple new ideas, knowledge growth tends to increase in scale. Allowing for a combinatorial explosion across the board, however, Weitzman's model fails to consider why knowledge grows in trajectories (Dosi 1982) and, in particular, why some trajectories grow faster than others (Nelson 2003). Furthermore, while in principle all ideas can be recombined with one another, there are numerous reasons why this does not actually happen. First, paraphrasing George Akerlof,[3] one may wonder if combining chicken and ice-cream is at all useful, or, said more generally, if all existing ideas can be recombined productively. Second, "[p]erceptions that certain technologies or components 'belong together' develop through social construction and previous association." For example, "if an electrical engineer of the 1940s had been asked about his profession's use of sand and aluminum, he probably would have replied with a blank stare. Today, he or she probably would reply that they are the most common basic materials of semiconductors and the focus of much research investment" (Fleming 2001, p.119). Third, the body of human knowledge is so large and dispersed (Hayek 1945) that only small portions of the potentially productive recombinations can be envisioned by boundedly rational actors, and only even smaller portions can be materialized. Indeed, the creative recombination of ideas is a problem-solving endeavour characterized by limited information-processing capacity (Simon 1977), and is based on a highly incomplete and uncertain search (Fleming 2001). Thus unlike Weitzman's model, the actual processes of knowledge recombination through which new knowledge is generated can unfold in very different patterns across technology domains (or, for that matter, across individuals, firms, institutions, etc.) as well as over time. In line with this view, the goal of this paper is to investigate how different dynamics of knowledge specialization and knowledge brokerage impact the rate of progress of technology domains.

We believe that focusing on technology domains provides a particularly useful perspective to study the growth of public technological knowledge. One notable advantage is that one can simultaneously address the rate and the direction of knowledge growth, which would hardly be feasible when looking at individuals or organizations (see Dosi 1982)[4]; while the variance observed across technology domains informs us about the mechanisms affecting knowledge growth in general, the direction of knowledge growth is determined by the growth rate of individual domains. Of course, paraphrasing Simon (1991), one may say that all inventions take place in the minds of individuals. As Simon was well aware, though, this does not mean that inventions take place in *isolated* individuals. To the contrary, inventions and inventors always build on existing technological knowledge and, in all domains of technology, the body of received knowledge "is the necessary background against which new insight emerges" (Staudenmaier 1985, p. 65). Thus, even the inventiveness of the most extraordinary genius and the most creative organization are deeply shaped by the state and evolution of the knowledge domain to which they are contributing, and this is true not just for technology (Rosenkopf and Tushman, 1998; Tushman and Rosenkopf 1992), but also for areas where the solitary work of genius is generally regarded as quintessential, such as philosophy (Collins 1998), art (Becker 1982), and mathematics (Davis and Hersh 1980). Progress in a technology domain is thus the result of an inherently public process, and quite often it happens that an inventor builds on technological knowledge generated by someone he or she does not even know (Ziman 1967; Knorr-Cetina 1999). To study the growth of public technological knowledge, therefore, a vantage point may be gained by focussing on the network of knowledge recombination at the level of technology domains, even if this comes at the cost of remaining agnostic about other relevant units of analysis such as "organizations, individuals, or other combinations of actors" (Marquis and Davis 2005, p. 337). From this perspective, inventors and organizations are then seen as relatively denser clusters of ideas within a knowledge recombination network, possibly linking "…multiple communities with highly specialized technologies and knowledge domains" (Boland and Tenkasi 1995, p. 350; Brown and Duguid 2001)[5]. In the next section we explicate the view of a knowledge recombination network; we then flesh out the concepts of domains' specialization and brokerage on the basis of that network.

2. Recombinant knowledge growth as an evolving network of technology domains

To formally define and, subsequently, empirically measure knowledge specialization and knowledge brokerage at the level of technology domains, we use a network-analytic approach. Technology domains may be thought of as cross-sectional "slices" of technological trajectories (Nelson and Winter 1982). And indeed, the notion that technological knowledge grows along trajectories of accumulation implies that the bulk of inventive knowledge recombination takes place within distinct technology domains, very similar to scientific disciplines being the loci of scientific knowledge production. However, like scientific knowledge in one discipline often spawns innovations in another, inventions developed in a technology domain can serve as an input for inventions in other domains. Recombinant knowledge growth can thus be modelled as an evolving network of technology domains, where arcs (i.e., valued and directed ties) indicate patterns of inventive recombination both within and across domains. Figure 1 is an example of such a network which depicts a cross-cut of the growth process of a hypothetical stock of public technological knowledge within a given time window. This stock consists of three technology domains, A, B, and C, in which 30, 40, and 10 new inventions, respectively, have accumulated over the given time interval. The arcs point to the domains from which ideas are taken and are drawn in the direction of knowledge *search*. Knowledge "flows" in the opposite direction of the arrows, though, and in the literature on diffusion (Rogers 2003), arrows are in line with the flows[6]. Let us now focus on domain A for an illustration. The inventions generated in A resulted from the recombination of (parts of) inventions from A's own knowledge base 25 times, from B's 30 times, and from C's 20 times. Or, equivalently, knowledge *spilled over*[7] 25 times from past to current inventions in A, 30 times from past inventions in B to current inventions in A and 20 times from past inventions in C to current inventions in A. During the same time interval, the inventions generated in A worked as an input for inventions in B 40 times, i.e., ideas spilled over from A to B 40 times.

----- Here Figure 1 Example: Recombination patterns within and between technology domains -----

Formally, a network N_t at time interval t is a four-tuple, $N_t = \langle J_t\ ,\ L_t\ ,\ V_t,\ A_t\rangle$ that consists of a finite set of nodes, $J_t = \{i,...,k,q,...,j\}$, a finite set of arcs (i.e., directed ties) between the nodes, $L_t = \{l_{ik,t},\ \ldots,\ l_{qj,t}\}$, a function $V_t(.)$ mapping arcs onto pertaining arc values h (i.e., tie weights), and a function $A_t(.)$ mapping nodes onto node values. Nodes represent technology domains and their values represent knowledge output. The arc value h_{ij} represents the number of times that ideas belonging to the right-hand subscript node have been used in idea-combinations of the left-hand subscript node, and arc directions point to the nodes benefiting from the recombination.

2.1 A network view on domains specialization and brokerage

Based on this network representation of recombinant knowledge growth, we can now elaborate the concepts of knowledge specialization and knowledge brokerage at the level of a technology domain; we will formalize these concepts in the operationalization section. Intuitively, a technology domain is specialized insofar as it grows through the recombination of a homogeneous body of closely related knowledge. In contrast, a domain is brokering insofar as it grows by recombining ideas across a broad spectrum of mutually unrelated technological areas. To flesh out and to model this intuition, our starting point is a domain's *recombinant niche*, i.e., the sub-network comprising a focal domain, the domains from which it recombines ideas (the source domains), the valued and directed ties linking the focal domain to its source domains, and the valued and directed ties linking source domains among each other[8].

In general, the degree of brokerage in a focal node's niche indicates the extent to which the focal node brokers between otherwise disconnected nodes (Burt 1992). Applied to our network model, the brokerage of a domain's niche indicates a pattern of knowledge recombinations from domains that in their turn do not recombine knowledge from one another or from the focal domain. Accordingly, the degree of brokerage of a domain's recombinant niche indicates the extent to which its underlying technological community, in the aggregate, brings together knowledge across unrelated and, thus, heterogeneous source domains (Figure 2, Panel A). Conversely, a domain with high specialization recombines closely related knowledge, i.e., knowledge from either the focal domain itself, or from domains that are strongly related among each another and to the focal domain (Figure 2, Panel C). Figure 2 shows four hypothetical domains, ranging from low specialization (Panel A) to high specialization (Panel D).

--- Here Figure 2. From A to D, increasing specialization of focal technology domain (grey node) ----

3. Theory and hypotheses

Having posited specialization and brokerage as opposites on a continuum, we are left with the paradox that, according to extant theory and empirical evidence, both specialization and brokerage enhance knowledge growth. To resolve this paradox, and to specify a theory of knowledge growth where the effects of specialization and brokerage are explicitly related, it is necessary to distinguish between specialization (and, consequently, brokerage) as a *property* and as a *process*. Applied to a technology domain, specialization is a *property* indicating that, at a given point in time, ideas are recombined from a homogeneous body of closely related knowledge. For example, our data show that the technology domain of "mechanical guns and projectors" is highly specialized (like Panel D in Figure 1) while, in contrast, the domain of "coating processes" is highly brokering (like Panel A). Seen dynamically, specialization indicates the *process* by which ideas are recombined from a body of increasingly related knowledge. For example, the domain of "superconductor technology: Apparatus, material, process" became much more specialized between the beginning and the end of the 1990s. The property versus process distinction is not generally made in literature. As we aim to show in this section, however, revisiting knowledge brokerage and knowledge specialization in light of this analytical distinction helps to reconcile the two concepts in a unitary theory.

The central claim of the brokerage argument is that the more mutually unrelated are the ideas that one is exposed to, the more innovative (on average) are the ideas that one generates (e.g., Burt 2004).

The knowledge recombination perspective makes it possible to spell out the causal mechanism underlying this claim. The ideas that an actor (in our case, a technological community) is exposed to provide the inputs that the actor will consider in knowledge recombination. Thus, in general, the more one is exposed to ideas that are mutually unrelated, the more abundant and the more diverse are the potential new combinations that one sees and, therefore, the more numerous and innovative the ideas one generates. Supporting this view, Holyoak and Thagard (1995) have shown at the individual level that creative transfers of ideas occur by shifting mental models in a cross-fertilizing fashion, typically through analogies and metaphors. Regardless of the specific cognitive mechanisms involved, being confronted with ideas from diverse knowledge domains and applications prompts unexplored mental representations, thereby exposing potentially useful relations between previously unrelated ideas (Anderson and Thompson 1989; VanLehn and Jones 1993). A well-known example of the benefits of knowledge brokerage is Gutenberg's printing press, which resulted from the combination of ideas from the disparate bodies of knowledge that Gutenberg had studied, including metallurgy, press, ink, movable types, and the alphabet (Diamond 1997). Literature overflows with similar accounts of other inventions, both old and new (Mokyr 2002). More systematic empirical evidence is reported by Dunbar (1996), among others; he showed that scientists working in teams with a greater diversity of scientific backgrounds tend to solve problems by conceiving more innovative solutions. Evidence is also provided by Burt (2004), who found that firm employees embedded in brokering networks generate more numerous and creative ideas.

The benefits of recombining knowledge from diverse domains extend beyond the individual, resulting in more creative artistic and academic teams (e.g. Guimerà, Uzzi, Spiro, and Amaral 2005) and organizations (e.g. Hargadon 2002). Notice that, at these aggregate levels, knowledge can also be recombined by pooling together members with distinct specializations. Thus, the beneficial effects of knowledge brokerage derive from the aggregate spectrum of ideas that a team, organization, or community draws from collectively, whether through individuals with knowledge-brokering competencies (such as Gutenberg) or through the pooling of knowledge from individuals with diverse specialties. Applying the brokerage argument to the level of technology domains therefore suggests that the amount and the novelty of potential recombinations available to a technological community depend on how heterogeneous is the stock of public knowledge residing within the community's recombinant niche (as captured by Figure 2). This leads us to our first hypothesis[9]:

> H1: The higher the degree of brokerage of a technology domain at a given point in time, the higher its subsequent growth rate; similarly, the higher the degree of specialization of a domain at a given point in time, the lower its subsequent growth rate.

If knowledge brokerage increases both the amount of recombinant opportunities and their novelty, it follows that knowledge specialization reduces both of them. What, then, are the benefits of knowledge specialization? Seen another way, what are the limits of knowledge brokerage? An obvious scope condition is that while knowledge brokerage and knowledge specialization are opposite modes of knowledge generation, their effects are to some extent complementary; indeed, the expected benefits of knowledge brokerage would vanish in the absence of knowledge specialization. That established, we contend that the relation between knowledge specialization and knowledge brokerage can be fully understood only from a dynamic perspective. While technological growth is enhanced by the abundant recombinant opportunities associated with a high degree of knowledge brokerage (brokerage as a property), the positive impact of knowledge specialization rests on the fact that recombinant opportunities are more efficiently exploited by increasing specialization (specialization as a process). Applied to the context of technology domains, when a domain grows through increased knowledge specialization, knowledge is recombined more efficiently, i.e., with lower marginal costs. In contrast, when a technology domain grows through increased brokerage (i.e., when a technology domain is in the process of de-specializing), the marginal costs of knowledge recombination are much higher.

To elaborate on our argument, let us first spell out the mechanisms underpinning the process of specialization; on the basis of that, we will then be able to explain why increasing specialization yields efficiency gains. As said, we define the specialization of a technology domain as the extent to which it recombines ideas from a homogeneous body of knowledge. Starting from a highly brokering niche (Figure 2, Panel A), the homogeneity, and thus specialization, of a domain can increase in four ways. First and foremost, a technological community may draw a progressively larger fraction of its total recombinant inputs from the domain's own knowledge base, thereby increasing the proportion of "self" recombinations. Second, a community may concentrate on fewer source domains, thereby progressively shrinking its recombinant niche. In both cases, specialization increases as a result of *exploitation*, i.e. by a community increasing the *depth* and reducing the *scope* of its knowledge recombination patterns (Katila and Ahuja 2002), consonant with theory elaboration and integration in scientific research programs (Wagner and Berger 1985; 1986).[10] Figure 1, however, shows that the homogeneity of a domain's recombinant niche may also increase in two indirect ways. The source domains may become progressively more homogeneous due to an increased recombination of ideas among one another; and, a source domain may become progressively more similar to the focal domain by recombining a larger proportion of its recombinant inputs from the focal domain.[11]

Having clarified the mechanisms underpinning the process of specialization, let us now turn to the efficiency gains associated with it. In short, our argument is that the more the progress of a technology domain is achieved through progressive specialization, the greater the efficiency gains associated with exploiting an increasingly familiar and homogeneous recombinant niche. By combining ideas from the same subset of the technological landscape in a path-dependent fashion, exploitation increases familiarity with these ideas (Fleming 2001). At the individual level, familiarity with a subject matter due to accumulated experience enables the individual to handle larger chunks of information, thereby facilitating knowledge recollection and application (Chase and Simon 1973), which helps him/her to search for, appreciate, and pursue potential recombinations (Simonton 2000; Walberg 1988). Furthermore, problems related to familiar knowledge can be more effectively and more efficiently decomposed into simpler sub-problems (Eisenhardt and Tabrizi, 1995). Familiarity with a new technology domain, however, can be achieved only at the price of high fixed learning costs (Hayes 1989; Simonton 1991), typically requiring a substantial amount of domain-specific tacit knowledge (Gavetti and Levinthal 2000) and many years of preparation even for talented individuals (Weisberg 1993). Therefore, there is a scale advantage to spreading fixed learning costs over a relatively larger knowledge output and "[i]t is only when one has hit the frontier of one's primary specialization, where new items of interest are hard to find, that it might be cheaper to learn items outside that specialty" (Postrel 2002, p.306).

The efficiency gains associated with the process of specialization are amplified at higher levels of aggregation, where knowledge is recombined by actors dispersed throughout an organization or a technological community. In these cases, familiarity with the recombinant niche also results in more widely shared "embedding circumstances" (for example, with regard to the technical jargon, instruments, and testing criteria used in the inventive process), and thus it is associated with more effective and efficient communication among members. Moreover, in the context of public knowledge, every new relation that is established among ideas and inventions within a recombinant niche effectively homogenizes the knowledge therein; this, in turn, progressively reduces the cognitive and technical distances among a community's recombinant inputs. In summary, our argument is that due to the efficiency gains associated with the exploitation of an increasingly homogeneous and familiar recombinant niche, knowledge growth tends to be faster when technology domains advance by progressive specialization. By the same token, these efficiency advantages are forgone every time knowledge is recombined in a de-specializing fashion, i.e., when a domain's recombinant niche becomes more heterogeneous and knowledge-brokering. Accordingly, domains' growth rates should be lower during those times when previously unrelated knowledge enters a domain's recombinant niche.

> H2: The more the degree of specialization of a domain increases during a given time period, the higher its growth rate in that period; similarly, the greater the increase of brokerage, the lower the growth rate.

In addition to efficiency losses, exploratory recombinations and knowledge brokerage entail higher unpredictability. Although both exploitation and exploration are uncertain processes that may or may not yield knowledge growth, exploitation rests on known uncertainties involved in recombining knowledge from a more familiar niche, whereas exploration is based on unknown uncertainties inherent in distant search (March 1991; Fleming 2001; Cohen and Aston-Jones 2005). Notably, exploitation reduces the probability of dead ends because failed recombination attempts indicate less successful parts of a familiar recombinant niche (Fleming 2001; Vincenti 1990). When one ventures into brokering previously unexplored domains, hardly any prior information is available within the focal community regarding fruitless combinations that should be avoided. Furthermore, brokering knowledge means de-embedding knowledge from one community and re-embedding it in another; this entails passing more arduous cognitive and cultural barriers (Brown and Duguid 2001), and it may trigger political intricacies and irrational factors whose effects are hard to predict (Latour 1987). When Edison invented the light bulb, he was accused of "the most airy ignorance of the fundamental principles of electricity and dynamics" (quoted by Hargadon 2002, p.57). However, in addition to a greater risk of failure, and consistent with both Edison's experience and Hypothesis 1, brokering knowledge through explorative recombinations is also more likely to generate unusually fruitful inventions due to the greater innovative potential inherent in recombining heterogeneous knowledge inputs. These arguments lead to our third hypothesis.

> H3: There is greater variance in growth rates among highly brokering technology domains than among highly specialized ones.

As stated by Hypothesis 2, the greater the increase in the degree of specialization of a technology domain during a given time interval, the higher we expect its growth rate to be (net of differences in the number and novelty of inputs available in the domain's recombinant niche). In line with Hypothesis 1, though, the number and novelty of inputs potentially available for recombination vary inversely with a domain's degree of specialization for three reasons. First, as experience accumulates, most of the recombinations in a domain's niche have already been tried and the wells run dry. Second, when a domain becomes more specialized, its recombinant niche becomes more homogeneous, and thus the recombinant inputs that remain are more likely to yield cumulative refinements of existing ideas than to yield breakthrough inventions. Third, being exposed to increasingly unambiguous and taken-for-granted methods and understandings may engender a habit of reproducing those methods and understandings at all costs (March 2005), and thus it may reduce actors' ability and willingness to search for path-breaking solutions (Levinthal and Rerup 2006). Therefore, the higher the level of specialization of a domain, the greater the likelihood that the efficiency gains associated with a further increase in specialization will be offset by a lack of creative inputs and seminal ideas. From these arguments it follows that the effects of *increasing* specialization are inversely proportional to a domain's *degree* of specialization.

> H4: The positive effect of increasing specialization on a domain's growth rate becomes less pronounced as domain's degree of specialization raises, eventually reaching a point where a further increase in specialization hampers the domain's growth rate.

Our first four hypotheses focused on the relation between domains' specialization and brokerage on the one hand, and the (variance in) domains' growth rates on the other hand. Specifying hypotheses where the dependent variable is a *change* variable allows us to test postulated causal relationships more directly and unambiguously than can be done through *level* variables (Hsiao 2003). However, our theoretical arguments can also be extended to predict a level variable – the knowledge output accumulated within a domain during a given time interval. To appreciate the difference between the two types of dependent variable, consider the following example. Our first four hypotheses attempt to explain why, between the early 1990s and the late 1990s, the knowledge output generated in the technology domain "internal-combustion engines" grew by one fourth and the knowledge output generated in the technology domain "mineral oils: processes and products" decreased by one fourth. However, our first four hypotheses are silent about the absolute size of domains' knowledge output in the late nineties, which for "internal-combustion engines" is about three times as large as for "mineral oils: processes and products".

Synthesizing H1, H2, and H4, the following argument can be postulated about the absolute size of domains' outputs. While a brokering domain has greater recombinant potential for future growth than a specialized domain, that potential is realized only to the extent that a process of increasing specialization takes place, i.e., to the extent that the heterogeneous inputs in its recombinant niche are progressively related to one another through exploitation-driven recombinations. By the same token, however, the more this process goes on, and thus the higher the degree of specialization reached by a domain, the fewer and less breakthrough are the potential recombinations left to sustain the domain's future advancement. Therefore, taken together, H1, H2, and H4 entail that (*i*) a domain characterized by low specialization (i.e., high brokerage) has not (yet) exploited its recombinant potential, while conversely (*ii*) a highly specialized (i.e., low brokerage) domain has already exhausted most of it. Thus (*iii*) on average, the absolute size of accumulated knowledge output should be largest when a domain is at intermediate levels of specialization (and brokerage), when its pertaining technological community can carry out abundant exploitative recombinations within a still sufficiently heterogeneous recombinant niche. These arguments lead us to our fifth and last hypothesis.

> H5: The absolute size of the knowledge output generated in a technology domain within a given time interval varies concavely with the domain's degree of specialization.

4. Data and operationalization

To test our theory, we have chosen what is probably the largest stock of public technological knowledge currently existing, a database that describes all of the patented inventions (of which there are over two million), and all of the citations between them (of which there are over 16 million), granted by the USPTO between 1975 and 1999 (Hall, Jaffe, and Trajtenberg 2001). We use the USPTO data to indicate (*i*) the nodes of our knowledge recombination network, i.e., technology domains, (*ii*) the ties connecting these domains, signalling knowledge recombination patterns, and (*iii*) the domains' growth rates. On the basis of this network, we will then operationalize our explanatory and response variables.

4.1 Technology domains

The USPTO has expert patent classifiers who examine the claims made in each application document. After the content of an application has been analyzed by experts in the pertaining field, the application is classified according to a set of well-specified criteria. According to the 1999 concordance scheme, the United States Patent Classification (USPC) featured 418 3-digit, or primary, classes of technological knowledge, and over 120,000 subclasses. For the nodes of our network we chose to use the former, for four reasons. First, primary classes correspond more closely to well-circumscribed technologies or industrial sectors, and are therefore more reliable and robust than other partitions (Henderson, Jaffe, and Trajtenberg 2005, p. 462). Second, while some patents contribute to more than one subject, patents are assigned to only one primary class based on their "main inventive content", i.e. their most important knowledge contribution as perceived by the patent examiner (Earls, Smith, Wolf, Saifer, Rishell, Russell, and Rademaker 1997). Because primary classes do not overlap, in contrast to subclasses, only the former can be unambiguously operationalized as nodes in a network. Third, patents are periodically reassigned to patent classes in a retrospective fashion to reflect the emergence of new technological domains or the disappearance of existing ones. Clearly, the more narrowly one defines the technology domains the shorter the time scale within which these structural changes occur. Within our observation period, reassignments have been extremely rare at the level of primary classes, while they have been in the thousands at the level of subclasses. This makes primary classes a preferable unit of analysis for our research purposes[12]. Forth, as said, there are slightly over 400 primary patent classes, which is a large yet manageable sample size.

Although the examiners' judgement is to some extent subjective, we believe that the combination of the USPC system with the examiners' expertise yields high levels of accuracy, reliability and inter-subjectivity. Moreover, there is an extensive body of literature that uses patent classes to indicate technology domains. For example, Powell and Snellman (2004) used patent classes to trace the changing importance of technological sectors over time. Similarly, patent classes have been used to measure technological proximity by, among many others, Almeida (1996), Jaffe and Trajtenberg (1999), Hicks and colleagues (Hicks, Breitzman, Olivastro, and Hamilton 2001) and Frost (2001). Also, in studies at the national level, patent classes have been frequently used to indicate technology domains, for instance to measure countries' technological specialization and technological advantage (Soete 1987; Patel and Pavitt 1987; Cantwell 1989; Patel and Vega 1999).

Furthermore, for our research purposes, whether the USPC technological classification is in part a social construction matters less than it may seem at first sight. Let us take the domain of "power plants" as an example. The bottom line of our argument is that inventors operating in different technological communities are exposed to different recombinant niches. Thus, our assumption is that inventors busy with power plant technologies will try to keep stride with and build upon the inventions pertaining to that domain. To do so, they will search the USPTO database through the USPC technological classification scheme, whether that reflects an entirely objective representation of the underlying domain or not. The result of that search will shape the recombination inputs to which they are exposed and thereby, according to our theory, the amount and novelty of new knowledge they are able to generate.

4.2. Knowledge recombination patterns

Patents cite earlier patents, pointing out the public knowledge inventions draw from, the so-called *prior art*. Thus, patent citations are indicative of the recombination process underlying the creation of an invention. Following upon Griliches' seminal work (1979), in the last few decades many scholars, especially in the field of applied economics, have exploited from this property of patent citations to investigate dynamics of knowledge recombination and spillover. In addition to the indirect validation provided by such large body of empirical work, Jaffe, Fogarty and Banks (1998) devised a validity test of patent citation indicators, concluding that patent citations are "a valid but noisy measure of technology spillover", which finding was later confirmed by Jaffe and Trajectenberg (2002). Certainly, patent citation data must be treated with caution. Alcacer and Gittelman (2004) used new data available since 2001, making possible to disentangle the patent citations made by inventors from those added by patent examiners, and concluded that taking individual patent citations as indicators of knowledge recombination yields a risk for both type I and type II errors. Despite these risks, however, overall there is ample evidence that patent citations are a useful indicator of knowledge recombination. Furthermore, unlike many prior studies, we do not focus on individual patent citations, but on aggregated patterns at the level of technology domain. As also Alcacer and Gittelman's study showed, at this level of aggregation the patent citations added by the patent examiners *do not* significantly differ from the ones originally inserted by the inventors (Alcacer and Gittelman 2004). Finally, it is again important to notice that we are interested in how recombinant niches, capturing the different exposition of technological communities to recombinant inputs, affects knowledge growth. From this perspective, patents' prior art provides us with the information we're after – i.e. the recombinant inputs that are most immediately visible to a technological community. Thus, paradoxically, when there is a difference between the actual recombination inputs used for an invention, and the ones inferred from the prior art of the patent document, it may be argued that for us the latter are more relevant than the former, given that they are the ones that are most likely to be retrieved (and thus recombined) by the inventors interested in the focal patent.

4.3. Network evolution

In our model, we represent USPC 3-digit technology domains as nodes and citations of patents in one domain by patents in other domains as arcs, where arc weights indicate the number of citations and arrows point into the direction of citations[13]. Based on all knowledge patented in the USA between 1975 and 1999, Figure 3 shows the network of knowledge recombinations between technology domains. The network is highly connected since, on average, a domain has at least one patent citation to more than half of the other domains; the highest arc values are of domains citing themselves. For clarity, the thickets created by the weakest five percent of the ties and by the reflexive ties have been left out of the picture; if they had been included, the representation would be too dense for readers to see any network at all.

----- Here Fig. 3. The network of patented technological knowledge production in the USA ------

For the remainder of our analyses, in order to capture the evolution of the network we partitioned the observation period into five-year intervals, following Podolny and colleagues (1996). Clearly, this truncation yields a bias if domains' citations of older patents systematically point to different technology domains than do more recent citations. We assessed the magnitude of this bias by a Quadratic Assignment Procedure (Krackhardt 1987, 1988), regressing the network based on the most recent five-year windows (1995-1999) on the network based on the whole twenty-five-year observation period (1975-1999). This analysis showed that the two network configurations are virtually identical, yielding a correlation coefficient as high as 0.999 ($p<0.001$), and indicating that a network representation based on a 5-year interval is virtually as unbiased as one based on the whole twenty-five year period. To avoid spurious relationships between the variables of interest, we then opted to model our knowledge recombination networks as a time series of five non-overlapping networks, for the following time intervals: 1975-1979, 1980-1984, 1985-1989, 1990-1994, and 1995-1999.

4.4 Domains' growth

For an invention to be patented, it must consist of knowledge that is new, non-trivial, and applicable. Therefore, a patent is by definition an idea that advances the stock of public technological knowledge. Accordingly, patent counts are generally regarded as a valuable proxy for measuring knowledge growth if the success, or impact, of each patent is taken into account (Grilliches 1990). If a patented invention consists of knowledge that is useful for the generation of subsequent inventions, it will be cited. In the words of Gittelman and Kogut (2003, p. 380), "because certain patents open richer technological veins, the subsequent advances in related technical knowledge encourage more innovative efforts in that area and, hence, more patents. These, in turn, cite the initial patents that opened this avenue of technological innovation. It is this feedback that carves a trace in the patent patterns." Accordingly, a widely used indicator of the impact of a patent on the advancement of knowledge is the number of citations it received (these are called *forward citations*, Griliches 1990). As an indirect validation that a patent's forward citations capture knowledge contribution, forward citations were found to be positively related to received royalties (Giummo 2003), to intangible assets after controlling for R&D expenditure (Hall *cum suis* 2005), to the value of a patent in the eyes of the patent holder (Harhoff et al. 1999), and to the social value of a patent (Trajtenberg 1990). Forward citations were also directly validated as a measure of knowledge contribution through surveys of inventors and experts by Albert *cum suis* (1991), and by Jaffe *cum suis* (2000). To the best of our knowledge, no published study in the large body of empirical research on the topic has disconfirmed the validity of this measure.

At the level of technology domains, we measure knowledge output by counting all the patents granted within a domain over a given time interval, weighed by the number of forward citations they received. Because it weighs each patent by the number of forward citations it received, our measure is consistent with the well-established notion that any quantification of knowledge output must reflect both the amount of new ideas generated and the extent to which these ideas gain recognition (Simonton 2000; Walberg 1998; Weisberg 1993; Fleming 2007). Call our knowledge output measure M_{it}, where i indicates all domains in our study population and $t = \{1975\text{-}1979, 1980\text{-}1984, 1985\text{-}1989, 1990\text{-}1994, 1995\text{-}1999\}$. To calculate a domain's growth rate, we then take the percentage difference between the domain's knowledge outputs in subsequent time intervals: $(M_{t+1} - M_t)/ M_{t,}$. The choice to confine the measure of domains' knowledge output within a five-year interval may engender an error, given that less than thirty percent of all citations are made to patents less than five years older than the citing patent. As a matter of fact, it takes fifty years to catch ninety percent of all citations received by a patent (Hall et al 2001), which is well beyond our observation period. To assess the magnitude of this error, we took all patents granted in each domain during the first five-year interval (1975-1979), and counted the citations they received during the same period. Then we counted the citations received by those same patents up to 1999, and we calculated the correlation between the two measures. Both Pearson's and Spearman's *rho* values were above 0.97 and highly significant ($p<0.001$), indicating that it is acceptable to measure domain's knowledge output, and thus domain's growth rate, on the basis of five-year intervals.

4.5 Domain's specialization and brokerage

To model for each domain i knowledge specialization and knowledge brokerage, as illustrated by Figure 2, we adapt Ron Burt's well-established network brokerage measure (Burt 2004, 1992). While in Burt's model brokerage is the opposite of *constraint*, in our model brokerage (B_i) is the opposite of specialization (S_i), which we define as $B_i = 1 - S_i$. The concise form of our specialization model is expressed by Equation (1); time indices are left out for ease of reading.

$$S_i = \Sigma_j (p_{ij} + \Sigma_q p_{iq} p_{qj})^2 \quad (1)$$

When i has no arcs at all, not even to itself, we leave S_i undefined because, in our conceptualization, a domain that carries out no recombination at all is neither specialized nor brokering. Let us start by fleshing out p_{ij} , which, like in Burt's model, indicates proportional tie strengths of i's direct contacts, here the proportion of ideas taken, and thus $0 \leq p_{ij}, \leq 1$. For our knowledge recombination model, a few changes are necessary with respect to Burt's model, of which by far the most important one is the inclusion of arcs of nodes to themselves (see Figure 1), by allowing for the possibility that $i = j$ (equation 5 below). Then there are three more issues to consider when $i \neq j$. First, our specialization model should capture well that if source domain j frequently takes ideas from focal domain i while i only rarely takes ideas from j, then i is not highly specialized in j (although i is constrained by j in Burt's model), and hence p_{ij} should in our model be small. Second, if i is highly specialized in ideas from j while j is highly specialized in ideas from i, then p_{ij} should approach 1. Third, if i takes ideas a given number of times from j and if j takes more ideas from i that to some extent funnel back to i, then i*'s* level of specialization is higher and p_{ij} should have a higher value, accordingly. To express these three requirements for $i \neq j$, we first define a term s_{ij} that is independent of j's specialization on i,

$$s_{ij} = \frac{h_{ij}}{\Sigma_k h_{ik}} . \quad (2)$$

The index variable k indicates all of the nodes in i's ego network, that is, all nodes that i draws from directly, including i itself. The arc value, h, is the number of times that ideas belonging to the right-hand subscript node are used in combinations of the left-hand subscript node. The same story can be told from j's point of view (again $i \neq j$),

$$s_{ji} = \frac{h_{ji}}{\Sigma_l h_{ji}} . \quad (3)$$

The index l denotes the nodes in the ego network of j, including j itself.[14] By combining (2) and (3), we can define p_{ij} in a way that meets our requirements,

$$p_{ij} = s_{ij} s_{ji} \quad (4)$$

If the focal domain has an arc to itself ($i = j$), we define $p_{ii} = s_{ii}$ (5)

For definitions of proportional tie values for indirect contacts with j via q (where q is also in i's ego network, and $q \neq i$ and $q \neq j$) we follow Burt (1992). [15] Corresponding to our informal description of the process of specialization (Figure 2), Model (1) and its satellite definitions of proportional tie strength make clear that the increased specialization of focal domain i can happen in four ways, sharpening our intuition about Figure 1. First, it can happen by an increase in the proportion of self-recombinations (p_{ii} increases). Second, when the focal domain increases its concentration on a limited number of source domains (s_{ij} increase). Third, when there are more strongly interrelated source domains (p_{qj} increase). Fourth, when source domains use proportionally more ideas from the focal domain (s_{ji} increase), which partly return to the focal domain later on (depending on a given value of s_{ij}). Finally, to measure how i's specialization changes between subsequent time intervals, we subtract i's prior level of specialization from i's current level of specialization. [16]

5. Analysis

5.1 Statistical methods

There are four sources of possible non-independence in our data. First, time-varying factors could affect the growth of all technology domains in a similar way; these may include macro-economic fluctuations, the rapid generic increase of both the number of patents and the number of citations (Hall et al. 2001), and changing practices among USPTO officers, as well as other factors. To model away these temporal effects, we use period dummy variables. Second, non-independence could also occur within subsets of units, as Hall *cum suis* (2001) have shown by pointing out similarities in patenting and citation patterns within six USPC macro-technological areas ("1-digit" classification). Again, we will model these effects by means of dummy variables. Third, as domains are interconnected by knowledge flows, non-independence may also yield network autocorrelation; that is, the growth of a technology domain may affect the growth of its contact domains. To account for this specific kind of non-independence, we adopt an established method in the social network literature and use a network disturbance model (Leenders 2002) [17]. Finally, the fourth kind of non-independence that could arise in the context of our study is related to the panel structure of our data. There may be unobserved heterogeneity across technology domains; thus, repeated observations within units are likely to be more similar than between units in our data. For example, certain technology domains may inherently have greater potential for growth than other domains, for reasons that are either unknown to or unobserved by the researcher. We exploit the panel structure of our data to account for this possible unobserved heterogeneity, using both a fixed-effects and a random-effects model to test our hypotheses[18].

5.2 Results

Our dependent variable – the domains' percentage growth rate – is distributed along a fairly well-behaved Gaussian curve, with very few outlying observations featuring exceptionally high values (i.e., values higher than ten times the population mean). As it turned out, all of the outliers correspond to very small technology domains, which could explain why their percentage growth rate is so high. To make sure that our estimates are not unduly influenced by a few peculiar cases, we chose to remove all observations in which a domain's growth rate between two subsequent time intervals was higher than ten times the population mean. This resulted in the removal of six observations. Post-estimation analyses of the residuals confirmed the appropriateness of this choice. Further, we removed all observations corresponding to technology domains that received no citations at all during a given time interval (again, corresponding to very small technology domains), because the concepts of recombinant niche, specialization, and brokerage are meaningless in those cases. Lastly, we removed from the analyses the technology domain called "miscellaneous," because it is merely a residual class in the USPC patent system. As a result of these choices, our sample decreased from 1672 observations to 1639 observations[19].

----- Here Table 1. Correlation matrix, means, and standard deviations ------

In Table 1, we report descriptive statistics and a correlation matrix for all of the variables used in our analysis. Table 2 shows the results of our statistical test, on which we focus now. Models 1 and 3 are baseline models within the fixed-effects and random-effects frameworks, respectively; Models 2 and 4 add a triplet of covariates representing our first three hypotheses to, respectively, Model 1 and Model 3. As mentioned above, our response variable is a domain's growth rate, i.e., a domain's percentage growth between time interval t and t_{+1}. Since most inventions are made within firms, we controlled for the (log of the) number of firms operating in a domain at t, which turns out to have a negligible effect on a domain's percentage growth in all models. Furthermore, some authors have argued that the proportion of backward patent citations firms make to their own patents is indicative of their ability to exploit their own inventions, which in turn may have repercussions on firms' knowledge strategies and investments (Hall et al. 2001). To make sure these firm-level dynamics do not affect our estimates of interest, we calculated for each domain during any time interval the average ratio of firm self-citations and used it as control. According to our analysis, firms' ability to exploit their own inventions does have a positive effect on domains' growth rates, but the effect is statistically significant only in Model 3. Our measure of specialization does not look at individual patents but rather at the aggregate recombination patterns occurring at the level of technology domains. As a consequence, our measure does not take into account whether a domain's degree of specialization is the result of a set of similarly specialized inventions or, in contrast, the result of a combination of highly brokering and highly specialized inventions. Of course, our specialization measure cannot be applied to the level of individual patents, because the patent-by-patent citation network is acyclic. Thus, to control for these patent-level differences, we computed Hall et al.'s indicator of patent originality (Hall et al. 2001), which measures the number of patent classes cited by a focal patent and therefore can be regarded as a simplified version of our brokerage measure at the patent level. On the basis of that, we then calculated for each domain during any time interval the coefficient of variation in patents' originality. The coefficient of variation indicates the extent to which individual patents deviate from the domain's mean, in terms of originality. Therefore, the higher a domain's coefficient of variation, the more that domain consists of a heterogeneous composition of highly brokering and highly specialized inventions; in contrast, the lower a domain's coefficient of variation, the more that domain is comprised of inventions with a similar degree of specialization. Our analyses show that the abovementioned variable has a positive effect on domains' growth rates, but the effect is statistically significant only once the hypothesized effects of specialization are accounted for. In line with common wisdom and previous analyses (Hall et al. 2001), we find a strong positive (albeit non-monotonic) relationship between time and a domain's growth; as a reference category, we used the first time interval in our observation period, 1975-1979. Furthermore, again in line with common wisdom, Models 3 and 4 show that the technology domains belonging to the areas of "computers & communications" and "drugs & medicals" (which includes biotech) have grown the fastest over the observation period, and have grown significantly faster than domains in our reference category, i.e. the miscellaneous category "others". In contrast, the domains "chemicals" and "mechanical technologies" have grown significantly more slowly than the reference category.

--- Here Table 2 Dynamic effects of specialization and brokerage on domains' percentage growth ---

In the last three rows of the table, we report the estimates pertaining to our first three hypotheses. All three hypotheses are statistically supported within both the fixed-effects and the random-effects specifications. As predicted by Hypothesis 1, specialized domains grow, on average, slower than brokering ones. Namely, according to our fixed-effects estimates, a difference of one standard deviation in the level of specialization of a technology domain results, on average, in an 18 percent decrease in the growth rate achieved over the subsequent five years. Conversely, as predicted by Hypothesis 2, the process of specialization is positively associated with growth. Thus, a technology domain that increases its specialization by one standard deviation over a given time period sees its growth rate increase during that period by 4 and a half percent on average. However, as predicted by Hypothesis 3, the positive effect of increasing specialization is reversed for high levels of specialization. A one standard deviation increase in specialization in a highly brokering technology domain ($S \cong 0.05$) yields a growth rate that is 10 percent larger than when no specialization occurs and, if the process of specialization is more rapid, this difference can be as high as 36 percent. Conversely, for a technology domain characterized by a relatively high degree of specialization, the effect of further specialization is altogether reversed. For example, for a level of specialization of $S \cong 0.45$, a one standard deviation increase in specialization leads on average to a 52 percent decrease in growth rates.

To test Hypothesis 5, we estimated both a fixed-effects (Model 5) and a random-effects model (Model 6) along the lines of Models 1 through 4, with the difference that the dependent variable is now a level variable – knowledge output accumulated in a domain within a given time interval – rather than a change variable. Hypothesis 5 stated that the volume of knowledge generated in a technology domain within a given time interval has an inverted U-shaped relationship with a domain's degree of specialization. To model this non-monotonic relationship, we jointly estimated the effects of domain's specialization and the effects of domain's squared specialization. In Table 3, we report the results of the test, according to which the inverted hypothesized U-shaped relationship between specialization and knowledge output is statistically highly significant[20]. Namely, the effects of specialization are initially associated with larger volumes of accumulated output but that they become negative after some point. Hence, the domains that accumulate the largest knowledge output within a five-year interval are neither extremely specialized nor extremely brokering, but rather hover around the middle (the maximum being reached when S = 0.6).

---- Here Table 3. Effects of specialization and brokerage on domains' knowledge output ----

Finally, to test Hypothesis 3, we compared the variance in growth rates exhibited over the observation period by highly brokering and highly specialized domains, respectively. To identify the two groups, we took the 10% of domains with the highest levels of brokerage and the 10 % of domains with the highest levels of specialization as measured during the first time interval. To account for possible fixed effects in the growth patterns of domains, we transformed the response variable in deviations from unit means; hence, the variance observed reflects the extent to which a domain's growth rate over time deviates from the domain's own mean. Our hypothesis is that the growth rates of domains that were highly brokering during the first time interval will exhibit greater variance over the observation period then those of highly specialized domains. Table 4 shows the results of our test. As we predicted, the variance in growth rates is larger for the group of brokering domains than it is for the group of specialized ones, and the difference is highly significant, according to Levene's test (1960). A glance at the confidence intervals for the means gives a sense of how differently the two modes of knowledge generation operate. Within a 95% confidence interval, the growth rate of a brokering domain is nearly twice as volatile as the growth rate of a specialized domain.

6. Conclusions

6.1. Contributions of the study

We believe that the present study yields three main contributions to the extant literature. First, it clears up the widely used notions of knowledge brokerage and knowledge specialization, showing more precisely what they are and how they relate to each another. As noted by Postrel (2002), task or labor specialization on the one hand and knowledge specialization on the other should not be confused. While task specialization may entail knowledge brokerage, and a brokering task may be based on highly specialized bits of knowledge, knowledge specialization and knowledge brokerage define opposite patterns of knowledge recombination. The focus of this study was on the generation of knowledge, regardless of the underlying organization of tasks. Therefore, we departed from the commonly used empirical strategy of looking at firm-, network-, or industry-level knowledge growth, and focused our attention directly on the growth of technology domains. In so doing, we were able to explicitly take into account that knowledge specialization and knowledge brokerage are opposite modes of recombinant knowledge growth, and explicate how their putative effects are dynamically intertwined. In contrast to earlier studies that treated specialization and brokerage as independent drivers of knowledge growth, we used a large set of longitudinal data to demonstrate that knowledge brokerage creates a recombination potential that can be efficiently exploited only by a process of increasing knowledge specialization. Moreover, we showed that increasing specialization enhances knowledge growth at a declining rate, and that therefore the process of specialization is either alternated by knowledge brokerage or it will ultimately lead to stagnation. The picture that emerges from our analyses is congruent with descriptive accounts of the evolution of science (Kuhn 1962), and of industries (Abernathy and Utterback 1978; Dosi 1982; Tushman and Anderson 1986; Utterback and Suarez 1993), where progress reportedly results from long periods of path-dependent, incremental refinements within a given research program or paradigm, sometimes alternated with path-breaking paradigm shifts.[21] Whether and to what extent these phenomena are driven by oscillating regimes of knowledge specialization and knowledge brokerage are intriguing questions, as well as opportunities for future research. While bibliographic records are more cumbersome to deal with than patents, they could nonetheless serve as a useful starting point in examining whether our arguments apply to the development of scientific fields (De Solla Price 1965).

From a normative standpoint, our findings are compatible with the view that to maximize knowledge growth, a balance must be struck between knowledge specialization and knowledge brokerage. Borrowing from March (2005, p. 9), one might say that knowledge specialization tends to increase a desire for the purification of existing ideas, yielding "exquisite barrenness," while knowledge brokerage entails the glorification of the newest recombinations, yielding shallow ideas produced in "cascades of triviality." In line with this view, we showed that, at any point in time, the technology domains that generate the most knowledge are neither too specialized nor too brokering. Our results also show that looking at knowledge growth from a static perspective makes little sense, though, and that any fixed position on the specialization-brokerage continuum is doomed to be sub-optimal. In the long run, the issue is not to find the most productive point along the specialization-brokerage continuum. Rather, it is to oscillate between knowledge brokerage (to generate new veins of productive recombinant inputs when the wells start to dry out) and knowledge specialization (to efficiently exploit those recombinant opportunities). While timely switching between these two opposite modes of recombinant knowledge growth is likely to be hard, our theory warns against convenient yet non-efficacious equilibria.

As a second contribution, the present study extends our understanding of how public technological knowledge accumulates. A distinguishing trait of the knowledge-based economy is that a large share of newly generated technological knowledge is a public good (Mokyr 2002). Most scholars agree that, since the Scientific Revolution, the rate of accumulation of public technological knowledge lies at the base of unprecedented yet sustained economic growth (Jones 2005). The mechanisms driving the accumulation of public technological knowledge, however, have hardly been studied and, to date, recombinant growth in the context of public knowledge is modelled as an unrealistically unconstrained combinatorial process (Weitzman 1996; 1998). In this paper, we took up the challenge to identify how specialization and brokerage affect the returns of knowledge recombination in the context of public technological knowledge. To this end, we studied how technological knowledge has accumulated across technology domains over a twenty-five-year period in what is arguably the largest source of public technological knowledge worldwide – the US Patent and Trademark Office. Our analysis showed that the process of recombinant growth driving the accumulation of public technological knowledge is far from combinatorially unconstrained. Rather, at any point in time, the rate of future accumulation of public technological knowledge is affected in important and predictable ways by the extent to which the body of accumulated prior knowledge in a technology domain is specialized. While other mechanisms are likely to influence the pace and direction of the accumulation of public technological knowledge, our results provide a relevant starting point to make current models of public knowledge growth both more realistic and more useful. Furthermore, our study explicates a framework and a methodology that can be used to investigate other explanatory mechanisms too.

The third main contribution of our paper is that it sheds new light on the widely acknowledged fact that the domains of human knowledge grow at widely different rates (Nelson 2003). Earlier studies have argued that these differences depend on "demand-side" forces (e.g., Schmookler 1966), as well as on supply-side factors (e.g., Rosenberg 1974, 1983) such as the strength of the link between scientific knowledge and practical know-how, and relatedly, the difficulty of doing precise, reliable, and generalizable experimentation (Nelson 2003). By explicating how the intertwined effects of knowledge specialization and knowledge brokerage affect the growth of technology domains, our study unveiled a previously unexplored mechanism that explains why certain domains of technological knowledge grow fast while others grow slowly or even stagnate. This theoretical advance may yield a practical contribution, as well. In the knowledge-based economy, investing at the right time in the most profitable technology domain(s) is of high strategic importance because knowledge production capabilities develop in an irreversible fashion (Nelson and Winter 1982; Kogut and Zander 1992). While successful strategic positioning in an evolving technological landscape is a key competitive advantage both for organizations (Stuart and Podolny 1996) and for countries (Nelson 1993), managers and policy makers are typically faced with great uncertainty when making these strategic decisions. To aid them in sailing the high seas of the knowledge economy, we showed that a specialization-brokerage analysis improves our ability to predict both the risks and the rates of return associated with investments in a given portfolio of knowledge domains.

6.1. Limitations and opportunities for future research

Our study has some noteworthy limitations, which in turn offer opportunities for future research. One limitation is that our analyses are based on a markedly macroscopic perspective. In order to focus on the dynamics of technological accumulation, we abstracted away from the actors that participate in and organize the inventive process. Although our approach allowed us to relate the aggregate dynamics of knowledge recombination to the aggregate outcomes in knowledge growth, we were unable to analyze in any detail how the individuals, teams, organizations, institutions, and inter-organizational networks comprising a technological community operate to make up such aggregate outcomes. For example, we could observe the specialization/brokerage of the aggregate body of knowledge accessed by a given technological community, but we could not observe whether there is "resource partitioning" (Carroll 1985) within technological communities, at what level (e.g., individual, firm, type of organization, etc.) that would take place, or what it might mean for the dynamics of knowledge recombination and growth. Similarly, while our study showed that the processes of specialization and brokerage are crucial to understand the growth of technology domains, we did not observe *how* these changes in the aggregate patterns of knowledge recombination of a technological community occur. For example, how is new knowledge brokered into a technology domain? Does it require "heavyweight" organizational mechanisms, such as moving personnel from the source to the application domain, or perhaps mergers and acquisitions? Or could it be done by moving knowledge only? And, what implications does the employment of these alternative mechanisms of knowledge transfer have for the probability of success of brokering acts? Can part of the large variance in performance we observed among brokering domains be explained by the difficulties inherent in the usage of these micro mechanisms, rather than by the difficulties of brokering knowledge? Currently, we do not have answers to these questions.

Our macroscopic study would greatly benefit from a more microscopic approach, and there exist possibilities to extend our theory to lower levels of analysis. For example, it is well-known that business groups tend to experience diminishing returns over time (Granovetter 2005). Our conjecture is that this phenomenon may in part result from a decreasing ability on the part of the participating organizations to generate novel ideas and technologies. Simply put, our reasoning is as follows. At the outset, individual organizations enter a business group with their own distinctive competencies and specializations, thereby pooling together a diversified stock of knowledge, generating recombinant potential through knowledge brokerage. As time passes, however, knowledge exchange relations tend to strengthen within the group, which on one hand fosters efficient knowledge generation but on the other hand progressively depletes the recombinant potential available to the business group. Hence, newly born business groups should generally yield more innovative ideas and technologies, and thus greater returns on average, than long-established business groups, especially when the latter feature a low entry rate. However, newly born business groups should also run greater risks of failure due the cognitive and communication difficulties associated with recombining heterogeneous technologies, competences, and knowledge trajectories. Therefore, it seems to us that useful insights could be gained if future studies were designed to explore the effects and dynamics of knowledge specialization and knowledge brokerage at the level of business groups, and more generally across units of analysis other than the technology domain.

The present study focussed on the growth of publicly accessible, codified knowledge. In so doing, we glossed over the role of tacit knowledge. By definition, tacit knowledge tends to remain private and it can only partially be transmitted beyond an inventor's proximate social network. Therefore, in and of itself, tacit knowledge contributes to the accumulation of public knowledge to a limited degree. This, however, does not mean that the generation and exchange of tacit knowledge is irrelevant for the progress of public knowledge. On the contrary, knowledge that is codified and publicly accessible is unlikely to be understood and effectively recombined unless it is complemented by related tacit and taken-for-granted notions (Gavetti and Levinthal 2000). Precisely because tacit and codified knowledge are complementary dimensions of the same phenomenon, rather than distinct phenomena (Polanyi 1958, 1967; Cowan et al. 2000), it seems reasonable to assume that the unobserved flows of tacit knowledge occurring within and between technology domains are to a large extent congruent with the observable flows of codified knowledge we studied. Nonetheless, it would be ideal to study directly the role of tacit knowledge in the accumulation of public knowledge. While this goal may be hard to achieve by means of a large-scale quantitative research design, more appropriate methodological approaches and analytical techniques may be borrowed (or, one may say, recombined) from the large body of work on organizational routines (Nelson and Winter 1982).

BIBLIOGRAPHY

Abernathy, William, and James M. Utterback 1978. Patterns of innovation in technology. *Technology Review* 80: 40-47.

Albert, M., D. Avery, F. Narin, and P. Mcallister. 1991. Direct validation of citation counts as indicators of industrially important patents. *Research Policy* 20: 251-259.

Alcácer, Juan, and Michelle Gittelman. 2004. How do you know what you know? The role of inventors and examiners in the generation of patent citations. Report, Stern School of Business.

Almeida, Paul. 1996. Knowledge sourcing by foreign multinationals: patent citation analysis in the US semiconductor industry. *Strategic Management Journal* 17: 155-165.

Anderson, John R., and Ross Thompson. 1989. Use of analogy in a production system architecture. *Similarity and Analogical Reasoning.* S. Vosniadou and A. Ortony, editors. Cambridge University Press.

Arrow, Kenneth J. 1962. Economic welfare and the allocation of resources for inventions. *The Rate and Direction of Inventive Activity: Economic and Social Factors*. R.R. Nelson, editor. Princeton University Press.

Basalla, George, 1988. *The Evolution of Technology*. Cambridge University Press, Cambridge.

Becker, Howard Saul. 1982. *Art worlds*. University of California Press.

Boland, Richard J., and Ramkrishnan V. Tenkasi. 1995. Perspective Making and Perspective Taking in Communities of Knowing. *Organization Science* 6: 350-372.

Brown, John Seely, and Duguid, Paul. 2001. Knowledge and Organization: A Social-Practice Perspective. *Organization Science* 12: 198-213

Burt, Ronald S. 1992*. Structural holes: the social structure of competition*. Harvard University Press.

--------- 2004. Structural holes and good ideas. *American Journal of Sociology* 110: 349-399.

Cantwell, John. 1989. *Technological innovation and multinational corporations*. Basil Blackwell.

Carroll, Glenn R. 1985. Concentration and specialization: Dynamics of niche width in populations of organizations. *American Journal of Sociology* 90: 1262-1283.

Chase, W.G., and Herbert A. Simon, 1973. The mind's eye in chess. *Visual information processing*. Chase, W.G., editor. Academic Press.

Cohen, Jonathan D. and Gary Aston-Jones. 2005. Decision amid uncertainty. *Nature* 436: 471–472.

Cohen, Wesley M., and Daniel A. Levinthal. 1990. Absorptive-Capacity: a new perspective on learning and innovation. *Administrative Science Quarterly* 35: 128-152.

Collins, Randall. 1998. *The Sociology of Philosophies: A Global Theory of Intellectual Change*. Belknap.

Cowan, Robin, Paul A. David, and Dominique Foray. 2000. The explicit economics of knowledge codification and tacitness. *Industrial and Corporate Change* 9: 211-253

Davis, Gerald F., and Christopher Marquis. 2005. Prospects for organization theory in the early twenty-first century: Institutional fields and mechanisms. *Organization Science* 16: 332-343.

Davis, Philip J., Ruben Hersh. 1982. *The mathematical experience*. Birkhauser.

De Solla Price, Derek J. 1965. Networks of scientific papers. *Science* 149: 510-515.

Diamond, Jared. 1997. *Guns, germs, and steel: the fates of human societies*. W.W. Norton.

-------- 2001. Unwritten knowledge. *Nature* 410: 521.

Dosi, Giovanni. 1982. Technological paradigms and technological trajectories. *Research Policy* 11: 147-162.

Dunbar, Kevin. 1996. How scientists think: online creativity and conceptual change in science. *Conceptual structures and processes: emergence, discovery and change.* T. B. Ward, S. M. Smith and S. Vaid, editors. APA Press.

Earls, E., H. Smith, L. Wolf, R. Saifer, E. Rishell, D. Russell, and C. Rademaker. 1997. *Examiner handbook to the U.S. Patent Classification System*. United States Patent and Trademark Office.

Eisenhardt, Kathleen M., and Behnam N. Tabrizi. 1995. Accelerating adaptive process: product innovation in the global computer industry. *Administrative Science Quarterly* 40: 84-110

Fleming, Lee. 2001. Recombinant uncertainty in technological search. *Management Science* 47: 117-132.

Foray, Dominique. 2006. *Economics of knowledge*. The MIT Press, Cambridge, MA.

Frost, Tony S. 2001. The geographic sources of foreign subsidiaries' innovations. *Strategic Management Journal* 21: 101-123.

Gavetti Giovanni, and Daniel Levinthal. 2000. Looking forward and looking backward: cognitive and experiential search. *Administrative Science Quarterly* 45 113–137.

Gilfillan, S.C. 1935. *The Sociology of Innovation*. Follett Publishing.

Gittelman, Michelle, and Bruce Kogut. 2003. Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns. *Management Science* 49 366–382.

Granovetter, Mark. 2005. Business groups and social organization. *The handbook of economic sociology*. N.J. Smelser, and R. Swedberg, editors. Princeton University Press.

Griliches, Zvi. 1979. Issues in assessing the contribution of research and development to productivity growth. *Bell Journal of Economics* 10: 92-116.

---------- 1990. Patent statistics as economic indicators: a survey. *Journal of Economic Literature* 28: 1661-1707.

Guimerà, Roger, Brian Uzzi, Jarrett Spiro, and Luis A. Nunes Amaral. 2005. Team assembly mechanisms determine collaboration network structure and team performance. *Science* 308: 697-702.

Hall, Bronwyn, Adam B. Jaffe, and Manuel Trajtenberg. 2001. The NBER patent citations data file: lessons, insights and methodological tools. *NBER* Working Paper 8498.

--------- 2005. Market value and patent citations. *RAND Journal of Economics* 36: 16-38.

Hargadon, Andrew B. 2002. Brokering knowledge: linking learning and innovation. *Research in organizational behaviour* 24: 41-85.

Harhoff, Dietmar, Francis Narin, F.M. Scherer, and Katrin Vopel 1999. Citation frequency and the value of patented inventions. *Review of Economics and Statistics* 81: 511-515.

Hausman, Jerry A. 1978. Specification tests in econometrics. *Econometrica* 46: 1251 - 1271.

Hayek, Friedrich von 1945. The use of knowledge in society. *American Economic Review* 35: 519-30.

Hayes, John R. 1989. *The complete problem solver*. Erlbaum.

Henderson, Rebecca, Adam B. Jaffe, and Manuel Trajtenberg. 2005. Patent citations and the geography of knowledge spillovers: a reassessment: comment. *American Economic Review* 95: 450-460.

Hicks, Diana, Tony Breitzman, Dominic Olivastro, and Kimberly Hamilton 2001. The changing composition of innovative activity in the US: a portrait based on patent analysis. *Research Policy* 30: 681-704.

Holyoak, Keith. J., and Paul Thagard. 1995. *Mental leaps: analogy in creative thought*. MIT Press.

Hsiao, Cheng 2003. *Analysis of panel data*. Cambridge University Press, Cambridge.

Jaffe, Adam. B., Michael S. Fogarty, and Bruce A. Banks. 1998. Evidence from patents and patent citations on the impact of NASA and other Federal labs on commercial innovation. *Journal of Industrial Economics* 46: 183-205.

Jaffe, Adam B., and Manuel Trajtenberg. 1999. International Knowledge Flows: Evidence from Patent Citations. *Economics of Innovation and New Technology* 8: 105-136.

---------- 2002. *Patents, citations, and innovations: a window on the knowledge economy*. MIT Press, Cambridge, MA.

Jaffe, Adam B., Manuel Trajtenberg, and Michael S. Fogarty. 2000. Knowledge spillovers and patent citations: evidence from a survey of inventors. *American Economic Review* 90: 215-218.

Jones, Charles I., 2005. Growth and ideas. Ch 16. *Handbook of Economic Growth*. P. Aghion and S. Durlauf, editors. Elsevier.

Katila Riitta, and Gautam Ahuja. 2002. Something old, something new: A longitudinal study of search behavior and new product introductions. *Academy of Management Journal* 45: 1183-1194.

Kogut, Bruce, and Udo Zander. 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science* 3: 383-397.

Knorr-Cetina, Karin. 1999. *Epistemic Cultures: How the sciences make knowledge*. Harvard University Press.

Krackhardt, David. 1987. QAP Partialling as a Test of Spuriousness. *Social Networks* 9:171-86.

------------ 1988. Predicting with Networks: Nonparametric Multiple Regression Analysis of Dyadic Data. *Social Networks* 10:359-81.

Kuhn, Thomas. 1962. *The structure of scientific revolutions*. University of Chicago Press.

Latour, Bruno 1987. *Science in Action: How to follow scientists and engineers through society.* Harvard University Press.

Leenders, Roger T. A. J. 2002. Modeling social influence through network autocorrelation: constructing the weight matrix. *Social Networks* 24: 21-47.

Levene, H. 1960. Robust tests for equality of variances. *Contributions to Probability and Statistics*. I. Olkin, editor. Stanford University Press.

Levinthal, Daniel, and Claus Rerup. 2006. Crossing an apparent chasm: Bridging mindful and less mindful perspectives on organizational learning. *Organization Science* 17: 502-513.

March, James. 1991. Exploration and exploitation in organizational learning. *Organization Science* 2: 71-87.

-------------- 2005. Parochialism in the evolution of a research community: The case of organization studies. *Management and Organization Review* 1: 5-22.

Marshall, Alfred. 1936, 8th ed. *Principles of economics*. Macmillan and co.

Mokyr, Joel. 2002. *The gifts of Athena: historical origins of the knowledge economy*. Princeton University Press.

Nelson, Richard, editor. 1993. *National Innovation Systems: A Comparative Analysis*. Oxford University Press.

------------- 2003. On the Uneven Evolution of Human Know-How. *Research Policy* 32: 909-922.

Nelson, Richard and Sidney Winter. 1982. *An evolutionary theory of economic change*. Belknap Press.

Nolan, Patrick and Gerhard Lenski. 2006. *Human Societies: An Introduction to Macrosociology. Tenth Edition.* Paradigm Publishers.

Nonaka, Ikujiro. 1994. A dynamic theory of organizational knowledge creation. *Organization Science* 5: 14-37.

Patel, Pari, and Keith Pavitt. 1987. Is Western Europe losing the technological race? *Research Policy* 16 59-85.

Patel, Pari, and Modesto Vega. 1999. Patterns of internationalisation of corporate technology: location vs. home country advantages. *Research Policy* 28: 145-155.

Podolny, Joel M., and Toby E. Stuart 1995. A role based ecology of technological change. *American Journal of Sociology* 100: 1224-1260.

Podolny, Joel M., Toby E. Stuart, and Michael T. Hannan. 1996. Networks, knowledge, and niches: competition in the worldwide semiconductor industry, 1984-1991. *American Journal of Sociology* 102: 659-689.

Polanyi, Michael. 1958. *Personal Knowledge: Towards a Post-Critical Philosophy*. Routledge and Kegan

------------- 1967. *The Tacit Dimension*. Routledge and Kegan Paul.

Poelwijk, Frank. J., Daniel J. Kiviet, Daniel M. Weinreich, and Sander J. Tans. 2007. Empirical fitness landscapes reveal accessible evolutionary paths. *Nature* 445: 383-386.

Postrel, Steven. 2002. Islands of Shared Knowledge: Specialization and Mutual Understanding in Problem-Solving Teams. *Organization Science* 13: 303 - 320

Powell, Walter W., and Kaisa Snellman. 2004. The knowledge economy. *Annual Review of Sociology* 30: 199-220.

Rapoport, Anatol. 1955. What is information? *Synthese* 9: 157-173.

Rogers, Everett M. 2003, 5th edn. *The diffusion of innovations*. Free Press.

Romer, Paul 1993. Idea gaps and object gaps in economic development. *Journal of Monetary Economics* 32 543-573.

Rosenberg, Nathan. 1974. Science, Invention and Economic Growth. *Economic Journal* 84: 90-108.

------- 1983. *Inside Black Box: Technology and Economics*. Cambridge University Press.

Rosenkopf, Lori, and Michael L. Tushman. 1998. The coevolution of community networks and technology: Lessons from the flight simulation industry. *Industrial and Corporate Change* 7: 311-346.

Schmookler, Jacob. 1966. *Invention and economic growth*. Harvard University Press.

Schumpeter, Joseph. 1939. *Business Cycles*. McGraw-Hill.

Simon, Herbert A. 1977. Scientific discovery and the psychology of problem solving. *Models of discovery*. Reidel.

Simonton, D. Keith. 1991. Emergence and realization of genius - the lives and works of 120 classical composers. *Journal of Personality and Social Psychology* 61: 829-840.

------- 2000. Creativity - Cognitive, personal, developmental, and social aspects. *American Psychologist* 55: 151-158.

Smith, Adam. 1776. *An inquiry into the nature and causes of the wealth of nations*. Strahan & Cadell.

Soete, Luc. 1987. The impact of technological innovation on international trade patterns: the evidence reconsidered. *Research Policy* 16: 101-130.

Staudenmaier, John M. 1985. *Technology's storytellers: Reweaving the human fabric*. MIT Press.

Stuart, Toby E., and Joel M. Podolny 1996. Local search and the evolution of technological capabilities. *Strategic Management Journal* 17: 21-38.

Sutton, Robert I., and Andrew Hargadon. 1996. Brainstorming groups in context: effectiveness in a product design firm. *Administrative Science Quarterly* 41: 685-718.

Trajtenberg, Manuel. 1990. A Penny for Your Quotes: Patent Citations and the Value of Innovations. *The Rand Journal of Economics* 21: 172-187.

Tushman, Michael L., and Philip Anderson. 1986. Technological discontinuities and organizational environments. *Administrative Science Quarterly* 31: 439-465.

Tushman, Michael, and Lori Rosenkopf. 1992. On the organizational determinants of technological change: Toward a sociology of technological evolution. *Research in Organizational Behavior* 14: 311-47.

Usher, Abbott Payson. 1954, 2nd ed. *A history of mechanical inventions*. Harvard UP.

Utterback, James M., and Fernando F. Suárez. 1993. Innovation, competition, and industry structure. *Research Policy* 22: 1-21.

VanLehn Kurt, and Randolph M. Jones. 1993. Learning by Explaining Examples to Oneself: A Computational Model. *Foundations of Knowledge Acquisition: Cognitive Models of Complex Learning*. S. Chipman, and A.L. Meyrowitz, editors. Kluwer.

Vincenti, Walter G. 1990. *What Engineers Know and How They Know It: Analytical Studies from Aeronautical History*. John Hopkins University Press.

Von Tunzelmann, G. N. 1998. Localized technological search and multi-technology companies. *Economics of Innovation and New Technology* 6: 231-255.

Wagner, David G. and Joseph Berger. 1985. Do sociological theories grow? *American Journal of Sociology* 90: 697-728.

------------ 1986. Programs, theory, and metatheory. *American Journal of Sociology* 92: 168-182.

Walberg, H. J. 1988. Creativity and talent as learning. *The Nature of Creativity: Contemporary Psychological Perspectives*. R. J. Sternberg, editor. Cambridge University Press.

Walker, Amasa. 1867 3rd ed. *The science of wealth: a manual of political economy; embracing the laws of trade, currency, and finance*. Little Brown and Company.

Weisberg, Robert W. 1993. *Creativity: beyond the myth of genius*. Freeman and Company.

Weitzman, Martin L. 1996. Hybridizing growth theory. *American Economic Review* 86: 207-212.

--------- L. 1998. Recombinant growth. *Quarterly Journal of Economics* 63: 331-360.

Young, Allyn. 1928. Increasing returns and economic progress. *The Economic Journal* 152: 527-542.

Table 1. Correlation matrix, means, and standard deviations

	Mean	St. Dev	% growth	Size	1985-1989	1990-1994	1995-1999	Comp & Comm	Drgs & Med	Elec & Elec	Mech	Chem	Spec	Δ Spec	Spec * Δ Spec	Know. output	Var. patent orig	Log nr firms
Percentage growth	0.5	0.5	1.00															
Size	6.1	1.5	0.05	1.00														
1985 thru 1989			0.16	-0.06	1.00													
1990 thru 1994			-0.07	0.03	-0.33	1.00												
1995 thru 1999			0.10	0.08	-0.33	-0.33	1.00											
Computers & Communications			0.33	0.08	-0.00	-0.00	0.00	1.00										
Drugs & Medicals			0.18	0.10	0.00	0.00	-0.01	-0.06	1.00									
Electrical & Electronics			-0.04	0.09	-0.00	-0.00	0.00	-0.11	-0.07	1.00								
Mechanical			-0.11	-0.08	0.00	0.00	-0.00	-0.19	-0.12	-0.24	1.00							
Chemicals			-0.13	0.10	0.00	0.00	-0.00	-0.14	-0.09	-0.18	-0.30	1.00						
Specialization	0.3	0.2	-0.10	0.25	0.03	-0.03	-0.07	-0.07	0.06	-0.01	0.08	-0.22	1.00					
Δ Specialization	-0.0	0.1	0.15	0.02	-0.01	0.03	-0.05	0.05	0.00	0.01	-0.01	-0.04	-0.30	1.00				
Specialization * Δ Specialization	-0.0	0.0	0.05	0.24	-0.02	-0.00	0.02	0.01	-0.01	-0.02	-0.03	0.03	-0.16	0.41	1.00			
Knowledge output	4901.1	6454.8	0.08	0.63	-0.14	0.05	0.29	0.05	0.14	0.11	-0.10	0.07	0.07	-0.02	0.07	1.00		
Variance patent originality	0.9	0.3	-0.11	0.04	0.06	-0.09	-0.23	-0.14	0.07	-0.03	0.07	-0.16	0.68	-0.07	-0.00	-0.14	1.00	
Nr firms (log)	253.3	271.3	0.01	0.68	-0.09	0.04	0.12	-0.02	0.11	0.10	-0.05	0.07	-0.01	-0.01	0.08	0.83	-0.13	1.00

Table 2. Dynamic effects of specialization and brokerage on domains' percentage growth

Dep. variable: Domain's % growth	**Model 1 Fixed-effects**	**Model 2 Fixed-effects**	**Model 3 Random-effects**	**Model 4 Random-effects**
Intercept	0.132 (0.56)	-0.011 (-0.05)	0.276*** (2.90)	0.166 (0.16)
1985 thru 1989 (*Ref.*: 1980 thru 1984)	0.279*** (6.44)	0.323*** (7.44)	0.245*** (5.97)	0.313*** (7.48)
1990 thru 1994 (*Ref.*: 1980 thru 1984)	0.072 (1.48)	0.106** (2.12)	0.035 (0.80)	0.101** (2.30)
1995 thru 1999 (*Ref.*: 1980 thru 1984)	0.211*** (3.91)	0.257*** (4.57)	0.175*** (3.84)	0.265*** (5.65)
Computers & Communications (*Ref.*: Others)			0.617*** (11.88)	0.608*** (11.97)
Drugs & Medical (*Ref.*: Others)			0.574*** (7.63)	0.562*** (7.63)
Electronic (*Ref.*: Others)			-0.026 (-0.59)	-0.036 (-0.82)
Mechanical (*Ref.*: Others)			-0.073** (-2.12)	-0.071** (-2.09)
Chemicals (*Ref.*: Others)			-0.137*** (-3.42)	-0.137*** (-3.42)
Number of firms (log)	0.014 (0.31)	0.023 (0.50)	-0.003 (-0.29)	0.006 (0.53)
Avg. firm self-citation ratio	0.003 (1.47)	0.001 (0.47)	0.004*** (2.45)	0.001 (0.81)
Variation in patent originality	0.121 (1.45)	0.265*** (3.15)	0.014 (0.24)	0.114* (1.69)
Specialization		-1.133*** (-4.46)		-0.342*** (-2.75)
Δ Specialization		0.484*** (2.44)		0.967*** (6.18)
Specialization * Δ Specialization		-1.354*** (-2.89)		-1.006*** (-2.63)
Number of units	415	415	415	415
Periods	4	4	4	4
Number of obs.	1639	1639	1639	1639
R-Squared (within)	0.111	0.156		
Log likelihood			-1016.8	-985.4

Random effects models are estimated by Maximum Likelihood Estimation
t ratios in parenthesis.
*** = p < 0.01, ** = p < 0.05, * = p < 0.1

Table 3. Effects of specialization and brokerage on domains' knowledge output

Dep. variable: Domains' knowledge output	**Model 5** **Fixed-effects**	**Model 6** **Fixed-effects**	**Model 7** **Random-effects**	**Model 8** **Random-effects**
Intercept	7.046*** (91.52)	5.901*** (56.58)	6.506*** (41.36)	5.182*** (32.08)
1980 thru 1985 (*Ref*.: 1975 thru 1979)	0.263*** (6.53)	0.182*** (4.88)	0.276*** (6.91)	0.174*** (4.70)
1985 thru 1989 (*Ref*.: 1975 thru 1979)	0.731*** (16.61)	0.662*** (16.29)	0.745*** (17.14)	0.653*** (16.20)
1990 thru 1994 (*Ref*.: 1975 thru 1979)	1.102*** (23.19)	1.035*** (23.58)	1.118*** (23.85)	0.102*** (23.59)
1995 thru 1999 (*Ref*.: 1975 thru 1979)	1.556*** (30.26)	1.531*** (32.33)	1.574*** (31.08)	1.524*** (32.63)
Computers & Communications (*Ref*.: Others)			1.199*** (4.17)	1.279*** (4.87)
Drugs & Medical (*Ref*.: Others)			1.317*** (3.12)	1.187*** (3.08)
Electronic (*Ref*.: Others)			0.905*** (3.64)	0.961*** (4.24)
Mechanical (*Ref*.: Others)			0.195 (1.00)	0.175 (0.98)
Chemicals (*Ref*.: Others)			0.757*** (3.47)	0.979*** (4.91)
Avg. firm self-citation ratio	0.001 (0.56)	0.004*** (3.07)	0.000 (0.27)	0.005*** (3.42)
Variation in patent originality	0.145** (2.27)	-0.102 (-1.65)	0.157** (2.51)	-0.106* (-1.74)
Specialization		5.787*** (13.95)		6.713*** (16.60)
Specialization squared		-4.637*** (-9.15)		-5.820*** (-11.78)
Number of units	415	415	415	415
Periods	5	5	5	5
Number of obs.	2045	2045	2045	2045
R-Squared (within)	0.720	0.764		
Log likelihood			-1840.3786	- 1667.2877
Random effects models are estimated by Maximum Likelihood Estimation t ratios in parenthesis. *** = p < 0.001, ** = p < 0.01, * = p < 0.05				

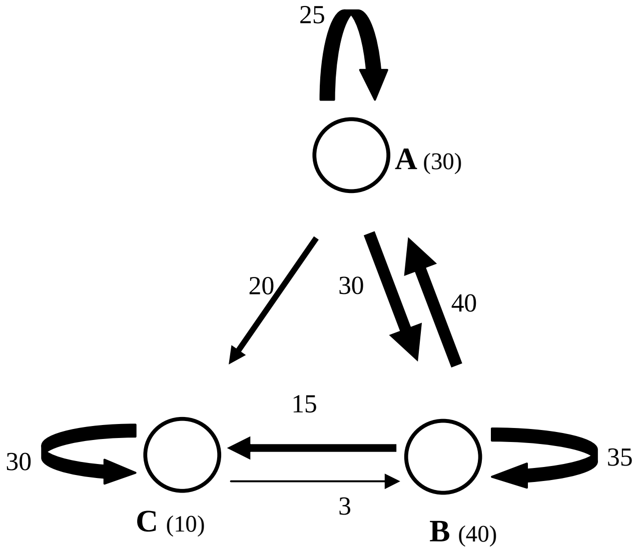

Figure 1 Example: Recombination patterns within and between technology domains

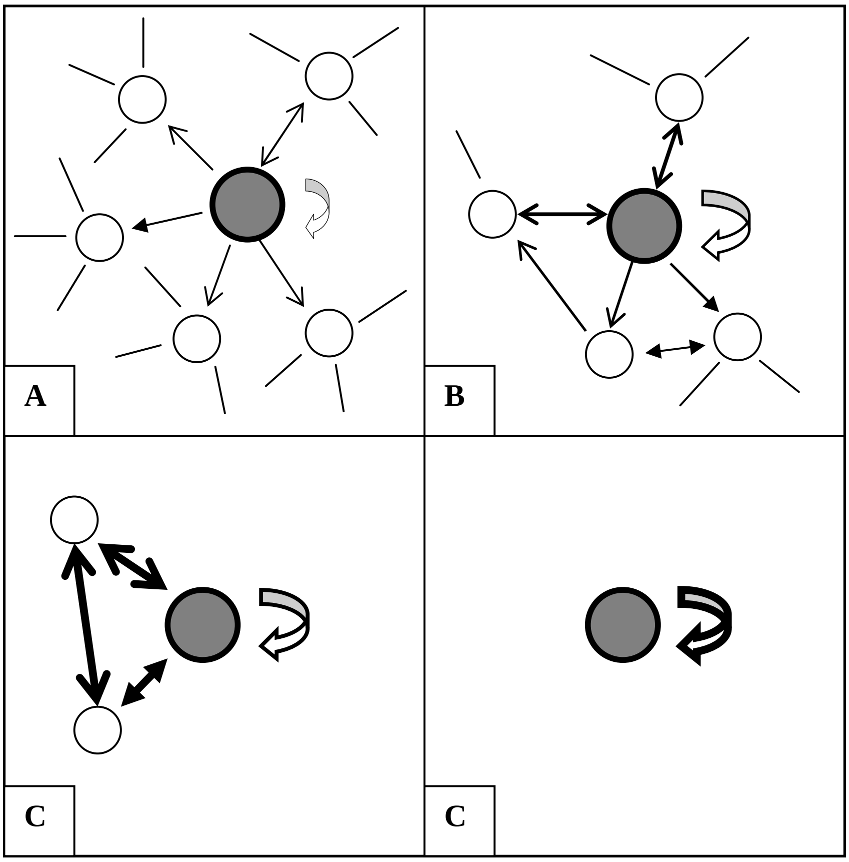

Figure 2. From A to D, increasing specialization of focal technology domain (grey node).

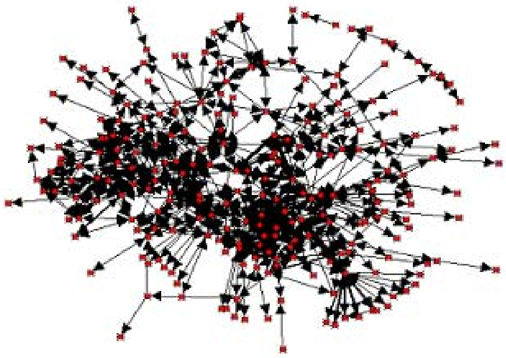

Fig. 3. The network of patented technological knowledge production in the USA.

ENDNOTES

[1] All economies are knowledge-dependent, but the notion of "knowledge-based economy" is used for those economies where knowledge is a more important input for production than capital, labor, and material resources, which holds for increasingly large parts of our modern economy. Knowledge is best comprehended dynamically, as a process of knowledge creation (or the result thereof): an entropy decreasing process of describing a part of the world more orderly (in the case of an innovation, a new part), i.e. with less information, in other words, by making it "more comprehensible with less effort" (Rapoport 1955: 171). This key point, of knowledge in terms of a difference between two information states, would be missed in a static definition of knowledge, for instance as symbolic representation, where knowledge becomes indistinguishable from information proper. An entropy-based definition of knowledge, at least in the context of our study, is consistent with the more traditional notion of "justified true belief" (e.g. Nonaka 1994), while it also connects to numerous other fields like communications science, physics, and biology.

[2] A few authors (e.g. Foray 2006) distinguish between knowledge generated by inventions as a result of knowledge recombination, and knowledge generated by discovering empirical regularities that were previously unknown. One may argue that even the latter mode of knowledge generation is ultimately combinatorial, because it involves synthesizing data and less general knowledge available at the time of discovery, some of which is embodied in scientific instruments. For that reason, DNA, for example, could simply not have been discovered in the Middle Ages. Whichever position one takes on this issue, one has to admit that to make sense of an observation, it has to be interpreted in terms of, i.e. combined with, existing knowledge, and that in general, knowledge recombination is undoubtedly an essential part of knowledge generation.

[3] Reported in Jones (2005), p.3

[4] Moreover, for a person or a firm, generating new knowledge can blur with other notions. For example, when a mathematics student learns the principles of calculus, either by exploring or exploiting, her personal knowledge grows while no new knowledge is created in society at large. At the level of technology domains, it is much easier to keep new and existing knowledge apart (see section on data and operationalization).

[5] Our conception of technology domains is therefore identical to the one adopted in the literature on epistemic communities. While technology domains are analytical categories that pertain to the partitioning of technological knowledge into technically-based categories, these analytic categories do not merely reflect objective properties of the knowledge. Rather, the categories defining technology domains co-evolve with the embedding circumstances prevailing within distinct networks of practice (Brown and Duguid 2001) and therefore with the cultural and institutional arrangements of distinct social (epistemic) communities (Knorr-Cetina 1999). Our focus differs from studies of epistemic communities in two ways, though. First, rather than investigating epistemic communities against the background of technology domains, we focus on technology domains and leave epistemic communities in the background. Second, rather than emphasizing how the categorical boundaries defining technology domains are negotiated and change over the course of time, we emphasize that once they become culturally and institutionally embedded, definitions of technology domains are fairly stable.

[6] This graphic difference between the literatures on diffusion and cognition (i.e. knowledge search) set aside, it is clear that knowledge recombination, be it brokerage or specialization, is a special case of knowledge diffusion, in this case of ideas from multiple sources to a target while being modified and recombined on their way. In fact the two literatures are not only consistent, but also perfectly complementary, one focusing on the generation of new knowledge and the other on the transmission of it.

[7] Knowledge spillover is a broader concept than knowledge recombination, as it also encompasses knowledge adoption and various other (e.g. unintentional) forms of diffusion. When used to describe the process of knowledge generation, as in the context of our study, an idea that spills over is one that becomes an input of recombination, contributing to the generation of a new idea. For an extensive discussion on spillovers and knowledge generation, see Jones (2004).

[8] This conception of niche as a relationally defined position in a network is similar to, and inspired by, the one employed at the level of the individual invention by Podolny and Stuart (1995).

[9] A mathematical representation of our hypotheses is reported in Appendix 1.

[10] Wagner and Berger's (1985, 1986) conceptualization of scientific research programs in terms of theory (1) elaboration and integration, (2) proliferation and variation, roughly matches our notions of (1) knowledge specialization, and (2) knowledge brokerage, respectively, which suggests a possibility of fruitful theoretical synthesis.

[11] Clearly, these direct and indirect mechanisms do not have an equally strong impact on a domain's specialization, and this should be accounted for. While in the methods section we will describe in detail how we weigh the impact

of each mechanism in our specialization model, for the time being it may suffice to know that, given the structure of our data, over 90% of a domain's specialization is captured by the first two mechanisms (i.e., by a community's own doing).

12 The very few primary classes which have either emerged or disappeared from the USPTO classification during our observation period are not reported in the NBER patent and patent citations database. Hence, consistent with our theoretical approach and empirical focus, our analyses are based on technology domains whose boundaries were stable and institutionally recognized during the whole observation period.

13 Note that as a consequence of aggregating patent citations, a domain can have a loop, or a reflexive tie, whereas an individual patent can't cite itself.

14 We checked the data for cases when $s_{ji} = 0$ while $s_{ij} > 0$, since one could argue that despite requirement (3) above, i would then still slightly specialize on j. That was hardly ever the case, leaving our results unaffected.

15 Thus $p_{iq} = (hi_{qi} + h_{qi})/\Sigma_k (h_{ik} + h_{ki})$ and $p_{qj} = (h_{qj} + h_{jq})/\Sigma_z (h_{qz} + h_{zq})$, but index variable z stands for nodes in q's ego network, including q, thereby allowing q to have reflexive ties while $i \neq j$, like in Burt's model. To check for programming errors, each of the authors computed exploitation with a different program and we correlated the results afterwards. Apart from rounding, the results were identical.

16 In the literature, the Herfindahl Index has been used to measure the degree of technological specialization of firms, industries, and countries, indicating the extent to which their patent citations are concentrated among patent classes (e.g., von Tunzelmann 1998). Our network-based conceptualization of specialization may be regarded as an extension of the Herfindahl index that weighs the concentration of patent citations across patent classes by two additional terms. The first term weighs the extent to which the cited patent classes cite one another, indicating similarity or homogeneity among a focal domain's source domains. The second term weighs the extent to which a cited class cites back the focal technology domain, indicating similarity or homogeneity among the focal and source domains. The Herfindahl index is recovered by Model (1) if we assume that the patent classes cited by a focal technology domain do not cite one another (i.e., if $p_{iq}\, p_{qj} = 0$), and that each source domain is fully specialized on the focal domain (i.e. $s_{ji} = 1$). In all other cases, Model (1) adjusts the specialization score to account both for the (dis)similarity of cited patent classes, and for the (dis)similarity of the focal domain with each source domain.

17 Network disturbance models exploit available information on the network structure of the data to account for the effects of possible network autocorrelation in the residuals. Let y be a $(n \times 1)$ vector of values of a network autocorrelated variable for n nodes making up a network. Further, let X denote a $(n \times k)$ matrix of values for the n nodes on k covariates. The claim that there is network autocorrelation in y implies that y_i is to some extent related to a weighted combination of y_j, where i and j are nodes, $j = 1…n$, and $i \neq j$. In the context of our study, this means that the growth rate of a focal domain is to some extent affected by a weighted combination of the growth rates of other domains. These weights are specified in a $(n \times n)$ matrix W based on the network data. Following a consolidated approach (Leenders 2002), we modelled the weight matrix W by row-normalizing our domain-to-domain patent citation matrix. Hence, in our study, W indicates for each domain what proportion of its total citations made, are made to each of the other domains. Within this framework, a network disturbance model accounts for the effects of network autocorrelation by modeling parameter ρ in the following statistical equation:

$$y = X\beta + \varepsilon, \qquad \varepsilon = \rho W\varepsilon + \upsilon, \qquad \upsilon \sim N(0, \sigma 2I)$$

As it appears, if W is well specified, the effects of network autocorrelation are removed from the residual, which is a necessary condition to obtain unbiased estimates of the parameters of interest.

18 Because it entirely removes between-unit variation (that is, it focuses solely on within-unit dynamics), the fixed-effects model provides a conservative approach yielding consistent estimates even in the presence of unobserved heterogeneity (Hsiao 2003). In contrast, based on a Hausman specification test (Hausman 1978), we cannot conclude that the estimates of the random-effects model are consistent in our case. Nonetheless, we chose to report the results of both the fixed-effects and the random-effects models because the latter makes it possible to estimate the effects of time-invariant control variables (something that is not possible within a fixed-effects framework). Therefore, in reading the results of our analyses, one should rely primarily on the fixed-effects model, while the random-effects model is reported to show that the estimates pertaining to our variables of interest are not significantly altered by the time-invariant controls that cannot be included in a fixed-effects model.

19 As a further warranty that our analyses are not unduly biased by the presence of very small domains, we also ran all of the analyses reported in the paper on samples based on technology domains that had, respectively, at least 50

and at least 100 patents granted within a given time interval. Although the sample size reduced to, respectively, 1530 and 1440 observations due to these thresholds, the estimations yielded qualitatively identical results to the ones reported in the paper. In fact, the only noticeable difference was that the estimates pertaining to our variables of interest had higher t-values and, hence, higher significance levels when we excluded the smallest domains. Furthermore, we have also collected data on the average R&D expenditure of technology domains. Because many R&D data were missing, we chose to report the models with the R&D control in Appendix 2, which is available at http://www.unisi.ch/print/personal-info?id=1529. Our estimates of interest were unaffected by the R&D control.

[20] In a model not reported in the paper, we also included a covariate for the number of firms patenting in each technology domain in any time interval. Although this covariate explained a very large portion of the variance in our dependent variable, the results pertaining to our hypothesized effects remained virtually identical. We chose not to include the abovementioned control variable in the paper because the number of firms patenting in a domain may depend on the growth of the domain, creating a potential endogeneity problem.

[21] Whereas these moves out of normal science were earlier seen as "punctuated equilibria," it has recently become clear that large shifts are less radical than was thought before, and that intermediate ideas are often polished out of history once their end results became clear. The steam engine, for example, came about by a series of inventions wherein each incrementally improved upon previous ones (Basalla 1988). In general, seemingly large jumps in a fitness landscape turn out to be evolutionary paths of intermediate steps that were previously overlooked (Poelwijk et al 2007).