VERY PRELIMINARY DRAFT, WHICH MAY BE SUBJECT TO SUBSTANTIAL CHANGES

PLEASE DO NOT CITE, QUOTE OR CIRCULATE WITHOUT PERMISSION

Technological Uncertainty in Markets for Technology

Ayfer Ali[1], Mariano Mastrogiorgio[2]

Abstract

Research on markets for technology has suggested that lack of participation in these markets generally stems from the need for complementary technologies in the licensing firm. Lack of participation may also be due to a lack of complementarity that unfolds over the time dimension. In fact, the value of a technological innovation to a firm may depend on the emergence of complementary technological innovations whose precise form and timing may be serendipitous and difficult to anticipate. This suggests that technological uncertainty may prevent firms from establishing a market-clearing price and may cause these markets to fail. In this paper we analyze the role played by technological uncertainty for the efficient functioning of markets for technology. For this purpose, a stylized theoretical framework is proposed, in order to shed light on a possible mechanism that may underlie the relationship between technological uncertainty and markets for technology. We argue that, in the presence of market incompleteness, technological uncertainty can make problematic the valuation of technological assets. Disagreements on valuation may then lead to less deals and, in general, to a less efficient functioning of markets for technology. In order to understand how technological uncertainty shapes the dynamics of commercialization, we explore a small market for technology whose actors are the technology licensing office (TLO) of a large academic medical center and the firms who showed interest by signing confidentiality agreements, options or licensing agreements for the TLO's patents. In order to measure the technological uncertainty of a patent and to assess its impact on the hazard rate of licensing, we adopt a novel methodology based on connectivity analysis, an approach originating from the field of network analysis and graph theory, which has found recent application in several studies on patent networks.

[1] Assistant professor, Department of Business Administration, Universidad Carlos III de Madrid
E-mail: aali@emp.uc3m.es

[2] PhD student, Department of Business Administration, Universidad Carlos III de Madrid
E-mail: mmastrog@emp.uc3m.es

1. Introduction

Recent literature on the topic of *markets for technology* has explored the issue of market efficiency from a perspective that merges the “market design” literature and the “markets for technology” literature (Gans and Stern 2010). The market design literature (Myerson 2008; Roth 2002, 2008) has highlighted three conditions for efficient market operations: market thickness, lack of congestion and market safety (Roth 2007). The markets for technology literature, on the other hand, has focused on the qualities that make technology and ideas different from more traditional goods: need for complementary ideas, value rivalry, and ease of reproducibility (Gans and Stern 2010). The issue of market existence and thickness is central in the debate and the lack of complementary ideas is likely “the most significant” of the issues that prevent thick market functioning (Gans and Stern 2010). Similarly, Ali and Cockburn (2012) have suggested that lack of demand in markets for technology is partly due to the necessity of complementary technologies in the licensing firm. This suggests that participation in markets for technology depends on whether these markets are able to include all the required complementary patent inputs, in other words on the ability of these markets to avoid strategic hold-up and to aggregate across a "package" of patents (Gans and Stern 2010). However, as suggested by Gans and Stern (2010), the lack of participation in these markets may also be due to a lack of complementarity that unfolds over the time dimension. In fact, the value of a technological innovation to a firm may depend on the emergence of complementary technological innovations whose precise form and timing may be serendipitous and difficult to anticipate (Gans and Stern 2010). Similarly, it may depend on the emergence of complementary technological innovations that may not yet exist (Rosenberg 1996). This suggests that technological uncertainty may play a problematic role for the functioning of markets for technology.

The aim of this paper is to analyze the role played by *technological uncertainty* for the efficient functioning of markets for technology. For this purpose, a stylized framework is proposed, in order to shed light on a possible mechanism that may underlie the relationship between technological uncertainty and markets for technology. We argue that, in the presence of market incompleteness, technological uncertainty can make problematic the valuation of technological assets. Disagreements on valuation may then lead to less deals and, in general, to a less efficient functioning of markets for technology. Therefore, the theoretical framework implicitly suggests that, in the presence of fundamental technological uncertainty, the achievement of market configurations characterized by more completeness can lead to a more efficient functioning. The achievement of configurations that tend to realize the ideal of market completeness seems to represent a feasible policy option, as demonstrated by the last developments in the IP industry, namely the birth of the IPXI, the first centralized financial exchange for technological assets. In order to understand how technological uncertainty shapes the dynamics of commercialization, we explore a small market for technology whose actors are the technology licensing office (TLO) of a large academic medical center and the firms who showed interest by signing confidentiality agreements, options or licensing agreements for the TLO's patents. In order to measure the technological uncertainty of a patent and to assess its impact on the hazard rate of licensing, we adopt a novel methodology based on connectivity analysis, an approach originating from the field of network analysis and graph theory, which has found recent application in several studies on patent networks (Barberà-Tomas et al. 2011; Martinelli 2012). Some preliminary evidence seems to suggest that higher levels of connectivity, and therefore lower levels of technological uncertainty, are correlated to a more than 11% increase in the hazard of licensing. However, we are very cautious about our results, because the sample size is small and several of our patent metrics may be correlated.

This paper aims to contribute to the literature on markets for technology, answering to the call of exploring the role played by uncertainty for their efficient functioning (Arora and Gambardella, 2010). It has been argued that the growth of these markets is hindered by uncertainty about the value of patents (Arora and Gambardella 2010). The first aim of this paper is to disentangle the problem of valuation, exploring the central role played by technological uncertainty. To our knowledge, previous research has not explored the linkage between markets for technology and technological uncertainty adopting this kind of perspective. The second aim of this paper is explore a novel methodological approach for the measurement of technological uncertainty.

The paper is organized as follows. In the next section a stylized framework is proposed, linking the uncertainty that unfolds in the technological domain to the problematic valuation of patent assets and the negative consequences for markets for technology. In the third and fourth sections the data and the empirical framework are described. In the fifth section some results are presented. The sixth section discusses some implications and concludes.

2. Theoretical Framework

Markets for technology and uncertainty about patent valuation

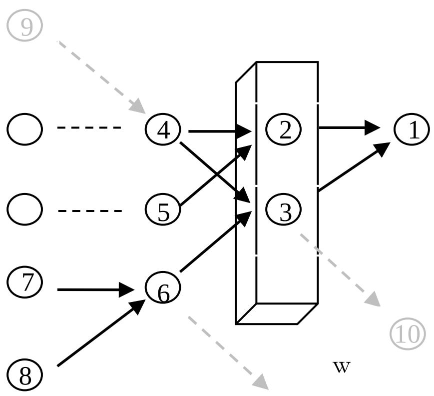
Even though the rate of transactions in markets for technology has been increasing, it is still uncommon for technologies to be traded in organized marketplaces (Gans and Stern 2010). As suggested by Fosfuri and Giarratana (2010), while most markets function "nicely and easily", markets for technology are plagued by "maladies" that can lead to their failure. Transaction costs have been explicitly recognized as a limit to the development of markets for technology (Arora and Gambardella 2010; Teece 1998). However, maybe due to a lack of definitional clarity, systematic research is still needed on the specific kinds of transaction costs and their influence on these markets (Arora and Gambardella 2010). Many models have stressed the problem of information asymmetry, accentuated by opportunistic behavior (Williamson 1973). However, these models have neglected the other pillar of transaction cost theory: the problem of uncertainty in the transactional environment, with especially severe consequences in the presence of bounded rationality (Arora and Gambardella, 2010). Few researchers have noticed this gap, starting to emphasize that "symmetric uncertainties" are more relevant obstacles to the functioning of these markets (Arora and Gambardella, 2010). In fact, it has been argued that the growth of markets for technology is hindered by uncertainty about the value of patents (Arora and Gambardella 2010). The value of patents is "skewed" (Scherer and Harhoff 2000) and, while firms may know the shape of this distribution, they may not know if the patent is in the "right tail" (Arora and Gambardella 2010). This suggests that the returns from buying or licensing in a technology are skewed: firms might therefore under-participate in markets for technology (Arora and Gambardella 2010). Indeed, as all IP professionals know, there are not only good deals and bad deals in markets for technology; there are also "no deals": using a sample of 229 US and Canadian licensors, Razgaitis (2004)'s survey has shown that, for a total of 100 licensable technologies, for only 25 of them a potential licensee is eventually found, negotiations are started in only 6 to 7 cases, and deals are eventually concluded in only 3 to 4 cases (Arora and Gambardella 2010). The survey has asked why deals were impeded, and the leading cause was in financial terms, because of disagreement on valuation.

Disentangling the uncertainty about patent valuation: the role of technological uncertainty

A fundamental prerequisite to understanding uncertainty about the value of patents is to provide a clear definition of the object of valuation: indeed it is very common to observe some confusion regarding the object of valuation, since in different circumstances it can mean the patent right alone, or the underlying technology alone, or the overall patented technology, where the latter comprises both the patent right and the underlying technology (Pitkethly 2007). It is therefore important to distinguish, at least conceptually, the value of the patent rights per se from the value of the underlying technology that the patent rights protect (Munari and Oriani 2011). Building on this distinction, we define two fundamental levels of uncertainty about the value of patents: 1) legal uncertainty, and 2) uncertainty about the underlying technological innovation (Munari and Oriani 2011). Before patent grant, legal uncertainty has two sources: uncertainty about whether the patent will be granted and, if granted, uncertainty about the scope of the claims (Lemley and Shapiro 2005). Once patents are granted, new sources of uncertainty arise, specifically as related to patent challenge and enforcement (Lemley and Shapiro 2005). While not the main topic of this paper, an example of the magnitude of such uncertainty can be helpful: when the U.S. Court of Appeals' decided to invalidate, two years before expiry, a patent on Prozac in 2000, the company´s stock price dropped 31 percent in one day (Lemley and Shapiro 2005). Uncertainty about the underlying technological innovation can be further decomposed into market uncertainty and technological uncertainty, as emphasized by Oriani and Sobrero (2008). Rosenberg (1996) identified and delineated a number of sources of technological uncertainty: namely, uncertainty about the development of complementary technologies (and about the development of broader technological systems), uncertainty about which technology will end up dominating the industry, as well as uncertainty about the possible uses of a technology. Uncertainty about uses deserves special attention if we consider that, very often, technologies come to the world with features and properties whose usefulness cannot be realized immediately, because of an inherent difficulty to foresee new uses well in advance (Rosenberg 1996). In fact, many technologies were developed to solve narrowly defined problems yet turned out to have significant and unanticipated applications in very different fields (Rosenberg 1996). An example of the inherent difficulty to foresee new uses is provided by the laser. Despite the significant impact of the laser on telecommunications, where it has revolutionized transmission, it has been reported that, initially, patent lawyers were even skeptical about applying for a patent, because according to them such invention had no possible relevance to that industry (Rosenberg 1996).The list of failures to anticipate technologies can be expanded without limits to prove that many technologies had very serendipitous life histories (Rosenberg 1996). Generally speaking, ex ante uncertainty about technological developments is “ontological”, because it doesn’t relate to just *whether* they will happen, it relates to *what* they will be, since they have never seen before (Lane and Maxfield 2005). As in the case of factors belonging to the legal domain, the channels through which technological factors may affect valuation is not straightforward. The relationship between technological uncertainty about uses and valuation, in particular, seems to be far more complicated than any superficial claim may suggest (Rosenberg 1996). Therefore, in the next sub-section, a stylized theoretical framework is illustrated in order to shed light on the complex relationship between technological uncertainty about uses and valuation. The theoretical model has been borrowed and adapted from the paper by Denrell, Fang and Winter (2003), in which they examine the challenges of imputing value to a resource without price guidance.

A stylized theoretical model, borrowed and adapted from Denrell, Fang and Winter (2003)

We define a “commodity resource” as a standardized asset for which many equal substitutable units exist. We then define a “complex resource” as an asset characterized by uniqueness and scarcity of substitutes, such as a patent. Commodity resources can be easily traded in an identifiable market. On the other hand, the market of complex resources, if it exists at all, can be highly imperfect (Dierickx and Cool 1989). Such imperfections make valuation problematic, and the problem of valuation is accentuated by uncertainties in the technological domain. To illustrate the problem of valuation in the presence of uncertainty, let’s consider a very stylized R&D process during which ideas/inventions, embedded in patents, are sequentially combined among each other into more complex ones that, at the end of the process, are combined into the design of a final consumer product. This is consistent with the view that technological development proceeds through path-dependent processes (David 1985; Dosi 1982) during which ideas and techniques are progressively accumulated or intersected (Levenhaghen et al. 1990). In the next figure, patents #4 and #5 are combined into patent #2, which is combined with patent #3 into the design of final consumer product #1, where patent #3 is the result of combining patents #6, #7 and #8. What is the value of a patent? Generally speaking, a proper valuation principle would be the one that "imputes" to the patent the returns that the patent makes possible through the product that it helps create.



In our stylized world, the value V_i of a patent i that can be directly combined into #1 is given by the revenue $f(S_1) - C_{i,1}$, where $f(S_1)$ is some function of the sales of the final consumer product in a market at a price P and $C_{i,1}$ is the cost of the research labor necessary to combine patent i into #1. The value of a patent that cannot be directly combined into #1 can be calculated by identifying the revenue that can be obtained by combining this patent into other patents that, in turn, can be combined into #1. However, patent #4 can be combined both into patent #2 and patent #3, that, in turn, can be combined into #1. The revenues that can be obtained through these alternative combinations have to be compared in order to identify the maximal revenue and therefore the value of patent #4. In general, to calculate V_i for a patent i that cannot be combined into #1 directly, we have to identify the maximum of $f(S_1) - \sum C_{k,j}$ among all possible *paths of use* by which patent i is combined into #1, where the k, j pairs define micro-paths of use like (2→1), (4→2) and (3→1), (4→3); both conceptually and computationally, we can formulate the whole problem within a dynamic programming framework (Bellman 1957), where the value V_i of any patent i has to satisfy the set of equations:

$$V_i = max_j\{V_j - C_{i,j}\}, \qquad i = 2,3, \dots ; j = 1,2,3, \dots$$

$$V_1 = f(S_1)$$

where the maximum is calculated over all possible patents j that patent i can be combined into (see Denrell et al. 2003). Given that we know the function of the sales of the final consumer product, we can use the recursive equation to find the values of all the patents that can be directly or indirectly combined into #1. Now, the central question is how the patent price forming in a hypothetical patent market relates to the patent value computed as above. If the market price were identical to the value computed as above, then the price of the patent would reflect the maximal revenue that the patent would make possible: in other words, *the price of the patent would precisely reflect the value of the patent in its best use among all possible uses*. It has been formally demonstrated that, in complete markets, prices coincide with the values as calculated above (Dorfman et al. 1958; Denrell et al. 2003). The condition of complete markets demands that each asset (patent) in the economy has a market and a price and, secondly, that each interaction among the economic agents is represented by some asset and, therefore, it is mediated by a market (Denrell et al. 2003). The principal implication is that, if the condition of complete markets is satisfied and therefore market prices correspond to the values as calculated above, then no knowledge about the set of all possible technological combinations would be necessary in order to identify the best way to use a patent. In fact, as emphasized repeatedly by Hayek

(1945), the principal claim for a system of prices is exactly that this type of knowledge is not necessary (Denrell et al. 2003). Rather, the owner of a patent would simply have to compare the values of $V_j - C_{i,j}$ for all possible patents j into which i can be *directly* combined: in few words, local and *decentralized* revenue comparisons would be enough in order to identify the best use of a patent (Koopmans 1957; Denrell et al. 2003). However, "*when markets are incomplete*[3]*, prices may not correspond to the values computed in the above way"* (Denrell et al. 2003, p. 983). For example, if patents #2 and #3 are not traded and therefore their prices are not observed, then the price of patent #4 cannot be expected to reflect the maximal revenue of combining patent #4 via #2 or #3 into #1, *unless economic agents know the possibility of these combinations*. In a few words, "valuation in incomplete markets depends crucially on the knowledge economic agents have about alternative transformations" (Denrell et al. 2003, p. 983). However, *in the technological domains, very often economic agents lack this kind of knowledge*. Knowledge of combinations and, generally speaking, knowledge *of uses* that are possible for a technology may be undermined by the simple fact that technological uses are often hidden and cannot be pre-specified and known ex-ante (Bonaccorsi 2011). In fact, as underlined by Basalla (1988), inventions offer a range of opportunities, but only few of them will be exploited during their life-time. We therefore argue that technological uncertainty about the uses of a technology, defined in terms of a systemic lack of knowledge of economic agents on its uses, may lead to valuation problems that, in turn, may cause the malfunctioning of incomplete markets for technology. Going back to the figure, if patents #2 and #3 are not traded and their prices are not observed, then the price of patent #4 cannot be expected to reflect the maximal revenue of combining patent #4 via #2 or #3 into #1, unless economic agents know the possibility of these combinations. However, if economic agents lack this knowledge, they cannot see "through the

[3] As in the case of markets for technology.

wall" (w) of uncertainty, facing fundamental valuation problems about patent #4, to the detriment of market agreements.

Proposition. Technological uncertainty about uses, making valuation problematic, hinders the functioning of markets for technology.

The relationship between technological uncertainty and markets for technology may vary across observable characteristics of a technology. In particular, the relationship may be moderated by another important source of uncertainty about the underlying technological innovation, namely how close is the technology to final commercial applications[4], both in terms of technological maturity as well as in terms of the extent to which it builds on early stage research (Narin et al. 1997; Ziedonis 2007). This insight holds other testable propositions. First, *technological maturity weakens the relationship between technological uncertainty about uses and the functioning of markets for technology*. As noticed by Ziedonis (2007), ''basic'' technologies seem to be characterized by higher uncertainty regarding the market potential, and this may negatively affect the decision to sign agreements that relate to the technology. In a similar vein, *the extent to which a technology builds on early-stage research strengthens the relationship between technological uncertainty about uses and the functioning of markets for technology*. However, this moderation role of early-stage research may become less intuitive if we consider that early-stage research may also play a positive role during the inventive process. As noticed by Fleming and Sorenson (2004), despite the variation across sectors in the degree to which inventions build on scientific research, several sectors seem to draw heavily on scientific research[5]. Within these sectors, the returns to applying science at the technology level may be influenced by the difficulty and the uncertainty of the inventive problem (Fleming and Sorenson

[4] That is, how close is patent #4 to patent #1 (see illustration).

[5] Such as the sectors of drugs and medicine. According to the study conducted by Mansfield (1995), 27% of the inventions of pharmaceutical firms required the application of science, compared to the 6% for electronics firms.

2004). In fact, as emphasized by Fleming and Sorenson (2004), "as the search space becomes increasingly complex […], local search routines break down, failing to identify the best combinations". Therefore, in these cases, scientific knowledge may represent the equivalent of a "map" that, reducing uncertainty, guides the search process towards a "directed identification of new *useful* [our italics] combinations" (pag. 910)[6], as the example of the discovery of the Prozac illustrates (Fleming and Sorenson 2004). This suggests that the extent to which a technology builds on early-stage research may eventually play a mutated role of moderation in the relationship between technological uncertainty about uses and market functioning.

3. Data

To test our propositions, we rely on an empirical setting based on US patent data. We explore a small market for technology. The market actors are the technology licensing office (TLO) of a large academic medical center and the firms who showed interest by signing confidentiality agreements, options or licensing agreements for the TLO's patents. Our dataset contains TLO's patents filed and granted from 1980 to 2008 and the associated agreements (confidentiality agreements, options, or licenses) signed with interested firms for those patents between 1980 and 2011. While a confidentiality agreement gives the firm the right to "look" a confidential description of the patent, an option (upon the payment of a fee) gives the firm the right to license the patent within a pre-specified period. Otherwise the "look" stage can be bypassed, and the firm can directly sign a license (for more details, see Ali and Cockburn 2012). Our objective is to isolate, at the patent level, the correlation between the technological uncertainty of a TLO's patent and the timing of licensing behavior. For this purpose, we build on

[6] However, the "know-what" elucidated by basic scientific knowledge may be bounded, as noticed by Barberà-Tomas and Consoli (2012), who have explored how "hybridization" (that is, the embodiment of competing operational principles within the same technology) may represent an alternative response, based on practical knowledge (know-how), to persistent uncertainty.

event history analysis in order to estimate the hazard rate of the time during which the first license occurs. As in Gans et al. (2008), the "failure" event is represented by the first instance of licensing. After trimming the dataset to exclude the few patents that were licensed before patent filing, our final dataset consists of 278 patents and of 2448 "time at risk" observations. In order to build our variables, we used patent data obtained from the USPTO and NBER databases (Hall et al. 2001), and from the Patent Network Dataverse database (Lai et al. 2011).

4. Empirical Framework

Econometric specification

In part we build on Gans, Hsu and Stern(2008) and we divide, for each TLO's patent, the data into yearly observations starting from the earliest filing date[7] and we define $\mathrm{license}_{\mathrm{it}}$ as equal to 0 until the year during which the first licensing event occurs for patent i, at which point a unique absorbing event sets $\mathrm{license}_{\mathrm{it}}$ equal to 1. We then define $\mathrm{post\,grant}_{\mathrm{it}}$, a time-varying regressor equal to 0 for years after the earliest filing date but before to the grant date, and equal to 1 for years after the grant date. This time-varying regressor allows us to discriminate between a pre-patent and a post-patent period for each patent (see Gans et al. 2008)[8]. We then introduce x_{it}, a time-varying regressor defined in the next section, which measures the technological uncertainty about uses of patent i. As in Gans et al. (2008), we employ a Cox Proportional Hazard Model, which includes a non-parametric baseline hazard rate, and a multiplicative term

[7] The "earliest filing date" is the filing date of the "parent" application, if the patent in question is a continuation or divisional application of an earlier US parent application.

[8] More importantly, according to Gans et el. (2008), since it is possible that a patent is licensed during the pre-patent period, the coefficient of *post-grant* may be interpreted as the "treatment effect" of patent grant on the timing of licensing. However, this interpretation is valid only under certain conditions, such as the "no anticipation of treatment" condition (Abbring and van den Berg 2003). Moreover, this interpretation builds on the "assumption that all selection effects can be captured by related observed and unobserved covariates" (Abbring and van den Berg 2003, p.1492). In general, this interpretation builds on some new literature on identification in duration models (for an overview, see Abbring and van den Berg 2003).

allowing time-varying and time-invarying regressors to have, relative to the baseline, a proportional impact (Lancaster 1990):

$$h(t, x_{it}, Z_i) = h(t) \cdot \exp\left(\beta_{x_{it}} x_{it} + \beta_Z Z_i\right) \cdot v_i$$

where $h(t, x_{it}, Z_i)$ is the hazard rate, at t, that license_{it} changes from 0 to 1[9], $h(t)$ is the baseline hazard rate, Z_i is a vector of controls, and v_i is a variable adding unobserved heterogeneity to the model. The term v_i captures the impact of variables that may act behind a spurious correlation between the timing of patent grant and the timing of licensing, such as a technology specific factor which cannot be observed by the econometrician (Gans et al. 2008). Ideally, we could introduce as many individual effects as the patents in our dataset. However, given the limitations in terms of the degrees of freedom that are necessary to estimate the parameters, we suppose that the distribution of v_i has a Gamma distributed functional form which can be summarized in terms of few frailty parameters (Jenkins 2005)[10] [11] at the level of the technological class.

Measures: technological uncertainty

The main independent variable is the technological uncertainty about the uses of a patent. We measure the technological uncertainty about the uses of a patent adopting a novel methodological approach based on the algorithmic count of the *paths of use* of which the patent is part. In order to identify the paths of technological use, we build on connectivity analysis, an

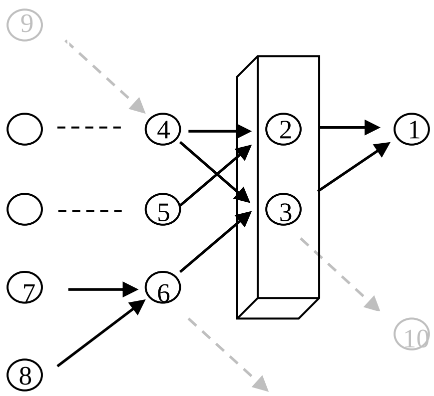
---

[9] That is, the instantaneous probability of failure at t, conditional on survival until t.

[10] If we define u = log(v), then the logarithm of the hazard rate with shared frailty is given by $\log[h(t, x_{it}, Z_i)] = \log[h(t)] + \beta_{x_{it}} x_{it} + \beta_Z Z_i + u$. As noticed by Jenkins (2005), we can think of this as a random effects model, which therefore assumes that the analyzed dataset consists of a hierarchal structure of different sub-populations whose differences relate to that hierarchy.

[11] At this point, some criticism could be raised about the decision to employ a continuous-time Cox Proportional Hazard Model when failure events, and therefore survival times, are eventually grouped into the same discrete time interval (i.e. a year). However it is important to emphasize that, when survival times are “tied”, several approximations can be used (often built in Stata by default, such as the Breslow approximation for tied failures) in order to derive the exact partial likelihood (Jenkins 2005). Generally speaking, historically many survival models assumed survival times to be realizations of a continuous random variable, and the application of discrete data to these models was not necessarily appropriate. However, as emphasized by Jenkins (2005), today this is much less of a problem, thanks to the increasing availability of methods for handling these data.

approach originating from the field of network analysis and graph theory, which has found recent application in several studies on patent networks (Barberà-Tomas et al. 2011; Martinelli 2012). Among the several approaches that have been proposed in connectivity analysis, we rely on the SPC (Search Path Count) approach, based on an algorithm that counts the uses of a patent by counting how many times a patent lies on all the possible paths between all patents of the citation network constituted by all US patents belonging to the technological classes that are relevant to the TLO's patenting activity (for technical details, see De Nooy et al. 2011). The connectivity approach is justified by the fact that, when we give a comprehensive look to patent citations, we are typically confronted with a network whose nodes and arcs are respectively constituted by patents and citations among patents, and in which using the number of the well known "forward citations" received by a patent amounts to use an in-degree centrality measure in order to assess the *local* importance of a node in the network (Valverde et al. 2006; Wartburg et al 2005). While the use of forward citations may be intuitive, "one may also suggest that this exercise ought to be integrated with a study of the *whole* [our italics] "connectivity structure" of the network in question", as suggested by Fontana et al. (2008, p.7). In other words, it is possible to conceive patent metrics that characterize the position of a patent within a citation network by taking into account not only direct citations (such as forward and backward citations), but also indirect citations. In this paper we pursue this approach, because we believe that direct citations are not sufficient to identify the paths of technological use and to reveal the "systemic" nature of uncertainty. As noticed by Wartburg et al. (2005), while in the case of social networks indirect links are less valuable than direct links, this is wrong in the case of patent networks, where we can hypothesize that the technological foundations of patents encompass not only the very recent developments that are directly cited, but also the developments provided by earlier patents. This is consistent with recent studies that have described the evolution of technology as a process characterized by "travels in time" that often resurrect early, or even extinct, technologies. The history of cornet bells is illustrating: after having classified a vast collection of cornet bells (some of them produced in 1825), Temkin and Eldredge (2007) showed that the evolution of cornet bells, and the evolution of technology in general, can be represented in terms of a "spreading, recursive network of pathways that often double back to "dead" ends" (Kelly 2010, p. 50). When we give a comprehensive look to patent citations, we are typically confronted with a network whose nodes and links are respectively constituted by patents and citations among patents. Firstly, such network is directed, since citations among patents have a direction, which is opposite to the direction of the knowledge flow among them. In the illustration below, patent #2 cites patents #4 and #5. We then assume a directed



flow of technical knowledge going from patents #4 and #5 to patent #2. The relevance of citations placed by patent examiners, rather than by inventors, may call into question the use of citations as maps of "knowledge flows", as noticed by Alcacer and Gittelman (2006). However, we have to emphasize that, given the nature of our research, the issue of knowledge flow as conceived by Alcacer and Gittelman (2006) doesn't matter very much as, in fact, we are more interested in the "technical link" among patents and not in weather the inventor was fully aware of the previous inventions (Martinelli 2010)[12]. Secondly, such network is binary, since the

[12] As noticed by Verspagen (2007), the use of patent citations as a tool for mapping the trajectories of technological knowledge can be justified by the fact that "a reference to a previous patent indicates that the knowledge in the latter patent was in some way useful for developing the new knowledge described in the citing patent" (p. 6).

presence of a citation can be associated to a 1 and its absence to a 0, and therefore no numerical weights besides 1 or 0 are associated to the citations so far. Thirdly, such network is not characterized by cycles. This property is intuitive, since patents can only cite previous patents. In the example, patent #2, issued in year 1990, cites patent #4 and #5 issued in 1985 and 1975 respectively. The presence of a cycle is excluded by the fact that patents #4 and #5 cannot cite a "future" patent such as #2. Fourthly, as noticed by Martinelli (2012), we can distinguish among three kinds of patents: a) startpoints, that is those patents with in-degree equal to 0, in which no arc is ending, such as #7, #8 and #9; b) endpoints, that is those patents with out-degree equal to 0, in which no arc is starting, such as #1 and #10; c) intermediates, that is those patents with in-degree and out-degree different from 0, such as #3, #4, #5 and #6. Once a patent citation network has been defined, the steps through which it has been built, and the measure of uncertainty has been obtained, are described below.

1. *Identification of all relevant technological classes*. For each of the 28 different technological classes of the patents licensed by the TLO (see Table 1), eventual overlapping technological classes were identified, and added to the set, relying on the classification information provided by the Uspto. The rationale of this decision is based on the fact that several technological classes of the US patent system overlap. For example, class 514 is considered to be an integral part of class 424, therefore retaining the same definitions. Starting from the technological classes of the patents licensed by the TLO, a set of almost 50 different technological classes was identified.
2. *Extraction of the time evolution of the citation network*. The time evolution of the citation network of all US patents belonging to these classes was extracted from the database of patent citations provided by the Uspto, containing several millions of records with all citations among all US patents of all classes issued between the beginning of 1975 and

the end of 2009. The time evolution of the citation network was extracted step by step, firstly extracting the citation network of patents issued in 1975 and adding to the citation network, year by year up to 2009, the patents issued in that year. In others words, 35 different "snapshots" of the final citation network were extracted: the first includes citations among patents issued in 1975, the second includes citations among patents issued between the beginning of 1975 and the end of 1976, the third includes citations among patents issued between the beginning of 1975 and the end of 1977, and so on up to 2009.

3. *Inversion of the citation network.* Each snapshot of the citation network was then inverted, since citations among patents have a direction which is opposite to the direction of the technical knowledge flow among them.
4. *Transformation of the inverted binary citation network in a weighted citation network.* Afterwards, for each snapshot, each citation in the network was assigned a numerical weight. The numerical weight was determined by a SPC (Search Path Count) algorithm. The SPC algorithm first identifies, for each startpoint (as defined above), all paths in the network between the startpoint and all the reachable endpoints. Then the algorithm counts, for each citation in the network, and therefore for the patents to which the citation is incident, the number of paths in the network passing through it. Therefore, such metric of "connectivity" measures the intensity of the technical knowledge flow passing through the patent. Going back to the illustration, the SPC algorithm firstly identifies all paths between startpoints #5, #7, #8 and #9 end endpoints #1 and #10: that is the 8 paths 9-4-2-1, 9-4-3-1, 9-4-3-10, 5-2-1, 7-6-3-1, 7-6-3-10, 8-6-3-1, 8-6-3-10. It then counts the

number of paths that pass through each patent. For example, 6 paths pass through patent 3 (for technical details, see De Nooy et al. 2011)[13].

5. *Re-scaling of patent metrics*. The connectivity of each patent was then re-scaled by the median connectivity of all patents of the snapshot network issued in the same year and belonging to the same technological class. The re-scaling was performed in order to remove systematic variation of connectivity values across technological classes, or taking place during time as the size of the network increases, as well as to purge the data of effects due to truncation[14]. We therefore followed the "fixed-approach" proposed by Hall et al. (2001) for forward citations, assuming that all sources of systematic variation are artifacts to be removed before comparing patents belonging to different technological classes as well as to different cohorts. The rationale of the decision to re-scale the connectivity of each patent by the median, rather than by the average, is the fact that the distribution of connectivity values is highly right skewed (see p. 25), and therefore very large values may easily distort averages (Herraiz et al. 2011). In order to correct for skewness, we took the logarithm of the scaled connectivity.

On the basis of the previous steps, we define the *scaled connectivity* of patent i at time interval t:

$$\mathrm{sc}_{\mathrm{it}} = \ln(\mathrm{c}_{\mathrm{it}}/\mathrm{Mc}_{(\mathrm{year,cl})\ \mathrm{t}})$$

where c_{it} is the connectivity of patent i at time interval t, and $\mathrm{Mc}_{(\mathrm{year,cl})\ \mathrm{t}}$ is the median connectivity at time interval t of patents issued in the same year and belonging to the same technological class. We then define x_{it}, a time-varying regressor which measures the

[13] The SPC method is based on en efficient algorithm that, after the topological sort of the network, computes the weights in fast time and exactly, that is without any kind of approximation (see Batagelj 2003).

[14] The issue of truncation shouldn't be particularly relevant in our case anyhow, since connectivity values have been calculated in a way that "fixed-window" comparisons are made possible (Hall et al. 2001).

technological uncertainty of patent i at time t in terms of decreasing levels of scaled connectivity.

Measures: control variables

Alternative sources of uncertainty about the underlying technological innovation. We control for alternative sources of technological uncertainty and for sources of uncertainty regarding the market potential of the technology, in order to rule out factors that may be correlated to our measure of technological uncertainty and, at the same time, to licensing outcomes. At the same time, we aim to check if the distance of the technology from market applications may play a moderating role in the relationship between technological uncertainty and licensing. As noticed by Ziedonis (2007), ideally, we could measure the technological uncertainty of a technology, as well as the uncertainty regarding its market potential, by determining the development's phase of the invention, and therefore its "maturity", at the time of disclosure. However, in our context, such detailed information is not fully available. Therefore, building on Lanjouw and Shankerman (2001) and Ziedonis (2007), we use the number of *backward citations* of the patent. As noticed by Ziedonis (2007), it is likely that, in technological areas with more prior art to cite, there is less technological uncertainty, as well as less uncertainty regard to the commercial potential of the technology under consideration. Extending the idea, we control for the aging of backward patent citations. We calculate the *average age of backward patent citations*, as the average difference between the issue year of the patent and the year of each backward patent citation. Similarly, we calculate the *minimum age of backward patent citations,* as the difference between the issue year of the patent and the year of the oldest backward patent citation. The years of backward patent citations were extracted through a computer algorithm from the USPTO web-page of each patent, since the USPTO database only provides backward citations dating back to 1975 at most[15]. We also control for sources of technological uncertainty that may be captured by the extent to which the patent builds on scientific research. We introduce the number of *non-patent references*, such as references to scientific journals, to control for the extent to which the patent builds on scientific research. As mentioned previously, the extent to which a patent builds on scientific research may determine the success of complex and uncertain inventive processes (Fleming and Sorenson 2004). At the same time, as noticed by Hegde (2011), the number of non-patent references can be used as a proxy of the closeness to commercial applications. In fact, as noticed by Narin et al. (1997), patents that tend to have more scientific references protect early-stage inventions. Extending the idea, we control for the *average age of non-patent references*, calculating the average difference between the issue year of the patent and the year of each non-patent reference, when available. We also control for the *minimum age of non-patent references,* calculating the difference between the issue year of the patent and the year of the oldest non-patent reference. Non-patent references and their years were extracted manually from the USPTO web-page of each patent[16]. Finally, we control for alternative sources of technological uncertainty that may be captured by the extent to which the technological "niche" to which the patent belongs grew over time before the patent was filed[17]. We introduce *growth rate of class*, a control for the growth rate of the number of patents issued in the same technological class of the patent between the two years preceding the earliest filing date of the patent.

Sources of legal uncertainty. As mentioned in the theoretical framework, there are several sources of legal uncertainty. And these may affect licensing outcomes. Before a patent is

[15] The algorithm was kindly provided by Juan Antonio Candiani, a colleague at Carlos III.

[16] The decision to opt for a manual extraction, rather than for an algorithmic one, has been justified by the fact that there is no consistent structure in the format of non-patent references on USPTO web-pages.

[17] That is when, supposedly, the invention was being conceived.

granted, there is uncertainty about whether the patent will be granted and, if granted, uncertainty about the scope of the claims (Lemley and Shapiro 2005). Once a patent is granted, patent grant and patent scope uncertainty are resolved (Gans et al. 2008)[18], even though other sources of pervasive uncertainty remain, specifically related to ultimate patent scope as well as patent challenge and enforcement (Lemley and Shapiro 2005). To control for sources of legal uncertainty, we introduce a *post-grant* time-varying regressor, equal to 0 for years after the earliest filing date but before to the patent grant date, and equal to 1 after patent issue (see Gans et al. 2008). Moreover, as in Gans et al. (2008), we introduce *grant lag*, a regressor that counts the number of years between the earliest filing date and the issue date[19].

Patent characteristics. We control for several patent characteristics, in order to rule out other factors that may affect licensing outcomes. The generality of a technology may play a relevant role. As noticed by Gambardella and Giarratana (2013), the capability of a firm to manufacture general-purpose technologies (Bresnahan and Trajtenberg 1995) represents an important determinant of licensing. In order to measure the generality of a patent, we count the *number of technological classes* according to the international patent classification (IPC), therefore adopting the measure introduced by Lerner (1994) as a proxy for patent scope. When counting the number of IPC classes, we use the first four digits only, as in Lerner (1994). Therefore, we count a patent assigned to IPC classes C12P 21/02, C12N 1/21, C12N 5/10, C07H 21/04 as falling into three classes, C12P, C12N and C07H respectively (Lerner 1994). We also control for

[18] Theoretically speaking, according to Gans et al. (2008) the key moment during which patent grant and scope uncertainty are resolved is when the "notice of allowance" is received by the inventor, rather than the patent grant date (which, on average, follows the allowance date after 5-7 months). Nevertheless, the results of their analysis seem to be robust to either the date of patent allowance or the date of grant as a key moment in which several sources of legal uncertainty are resolved.

[19] As mentioned in a previous footnote, according to Gans et al. (2008) the coefficient of *post-grant* may be interpreted as the treatment effect of patent grant on the timing of licensing. However, this interpretation is valid only under certain conditions, such as the "no anticipation of treatment" condition (Abbring and van den Berg 2003). Moreover, this interpretation builds on the "assumption that all selection effects can be captured by related observed and unobserved covariates" (Abbring and van den Berg 2003, p.1492). Gans et al. (2008) introduce a *grant lag* regressor (to be more precise, they introduce an *allowance lag* regressor) in order to directly control for the fact that *post grant* may be associated to an increase in the licensing hazard because of an underlying spurious correlation, due to potential unobserved features, between the *grant lag* and the *licensing lag* (where the *licensing lag* is the distance, in years, between the earliest filing date and the date of the first license).

the *number of claims*. Claims have been used as an alternative proxy of the generality of the patent (Gambardella and Giarratana 2013), following the idea that "the number of claims is […] an indication that an innovation is broader" (Lanjouw and Schankerman 2004, p. 448). At the same time, more claims may be an indication of higher scope of legal protection (Lanjouw and Schankerman 2004), that in turn may affect licensing outcomes. As noticed by Merges and Nelson (1990), claims define the boundaries and the legal basis of patent protection, forming a protective line around the patent that lets others know when they are infringing on their rights[20]. We also control for the "importance" of the patent. As noticed by Trajtenberg (1990), forward citations determine the importance of the patent and are correlated to the value of the underlying invention, and therefore may affect licensing outcomes. In fact, when a patent has many forward citations, that means that the patent is significant since many other inventors are building on it (Hall et al. 2001). However, the use of forward citations as a control may present some difficulties (see Mehta et al. 2009). More is the time that has passed since the patent was issued, more the citations accumulate. Therefore, more is the time that has elapsed between issue and the first license, more are the citations. In order to take this aspect into account, we introduce *forward metric*, a measure that scales the number of forward citations received by the patent until the year that precedes the first license (or, in the case of censoring, the year that precedes the last year of observation) by the average number of forward citations received, until the same year, by patents belonging to the same technological class and to the same cohort[21] (Ziedonis 2007). We also try to rule out factors that may be correlated to our measure of technological uncertainty and, at the same time, to licensing outcomes, that may be related to the experience of the inventors. As noticed by Fleming (2007), the distribution of inventive outcomes is highly

[20] The scope of legal protection can be abstractly defined in terms of a set of multiple "embodiments" (i.e. claims) of the technology that, analogous to the "metes and bounds" of a real property, distinguishes inventors' intellectual property from the surrounding terrain (Merges and Nelson 1990).
[21] The cohort being defined by the issue date.

skewed, and this should hold true if we measure inventive outcome in terms of patent citations, financial returns, or connectivity values in our case. Against the traditional belief that the outliers of this distribution (that is, innovative breakthroughs) arise from the effort of lone inventors, Singh and Fleming (2010) have demonstrated that collaboration between inventors increases the probability of breakthroughs, because of greater opportunity for recombination during the process of creative search. We therefore control for the *number of inventors* of the patent, as well as for the *number of inventors' patents* issued before the first license. Inventors' careers have been extracted from the Patent Network Dataverse database (Lai et al. 2011). In order to uniquely identify inventors' careers, we relied on the upper-bound disambiguation (Lai et al. 2011)[22].

Other controls. We introduce a *looked before* dummy, equal to 1 if the patent was "looked" before the first license.[23] We also control for the *number of times looked before* the first license. This last control may capture, in some extent, competition by other firms, that may have an impact on licensing outcomes. In fact, in a very similar setting, Ziedonis (2007) has introduced the variable "competitors", based on the count of the firms that signed secrecy agreements for a licensed patent.

Patent issue year fixed effects. In order to remove systematic sources of variation taking place during time and that, eventually, may not be captured by the rescaling of the connectivity metric, we introduce a dummy for the issue year of the patent.

[22] The lack of a consistent and unique identification of inventors at the USPTO results in name ambiguity on patent records. In order to remove ambiguity, a disambiguation algorithm has been proposed by Lai et al. (2011), leading to two clustering solutions: a lower-bound disambiguation, which attempts to capture the careers of inventors in their entirety, at the cost of lumping together, occasionally, different inventors; an upper-bound disambiguation, "which attempts to ensure that each cluster corresponds to a distinct inventor at the cost of occasionally splitting a single inventor over multiple clusters" (Lai et al. 2011, p. 19).

[23] That is, if a confidentiality agreement or an option was signed.

5. Empirical Results

Descriptive analysis

Before moving to the main results, the time evolution of the patent citation network and the statistical properties of patent connectivity are briefly described. The example in Figure 1 shows a "snapshot" of the time evolution of the patent citation network. As shown in Figure 2 and Table 2, the size of the network increases over time in terms of patents as well as in terms of citations, consistently with the broad empirical evidence (Hall et al. 2001). In Figure 3, the increase in the number of patents is decomposed in terms of start-points, end-points and intermediates. According to Martinelli (2012), an increase in the number of start-points, associated with a decrease in the number of end-points, it may be an indicator of the emergence of new streams of research and technological paths that converge to a limited set of paths, suggesting the presence of some kind of selection process, that in turn may reduce uncertainty. As shown in Figure 4 and 5, the time evolution of the network is marked by a sharp increase in the number of weak components[24] during the first years. This number peaks around 1980. It then decreases and stabilizes around the value of 2500 weak components, due to the emergence of a very large weak component that progressively occupies a larger portion of the entire network. The presence of a "giant component" can be observed in network configurations characterized by non-uniform distributions, such as power-law distributions[25] (Newman 2003). Networks

[24] In general, a component is a sub-network in which there is a path between all pairs of patents and there is no path between a patent in the component and a patent not in the component. In formal terms, a weak component is a maximal weakly connected sub-network. A network is weakly connected if each pair of vertices is connected by a semi-path, that is by a semi-walk in which no vertex between the first and the last vertex of the semi-walk occurs more than once. A semi-walk from vertex u to vertex v is a sequence of lines such that the end vertex of one line is the starting vertex of the next line and the sequence starts at vertex u and ends at vertex v (De Nooy et al. 2011).

[25] A distribution is power law if $p_k = ck^{-\alpha}$, where p_k is the probability that a node has degree k, and α is a scaling parameter. A network is characterized by the presence of a giant component when the condition $\sum_k k(k-2)p_k > 0$ is satisfied. The condition

characterized by power-laws, sometimes referred to as "scale free networks", have received increasing attention, since they have been observed in a series of real world structures, such as the WWW, the internet and, notably, patent citation networks. In the case of patent citation networks, Valverde et al. (2006) have found that the distribution of in-degrees (forward citations) obeys an extended power-law that, below a certain threshold of the in-degree, reduces to an exponential distribution. In the case of patent connectivity c_{it}, we expect to observe similar distributional properties. Figure 6 plots the distribution of patent connectivity on log-log scales, as it evolves over time. Figure 7 expands the plot of the distribution for the 1975-2009 period[26] [27]. The horizontal axis of each panel indicates the logarithm of patent connectivity, while the vertical axis indicates the logarithm of their cumulative probability distribution. As it can be noticed, the distribution is relevantly right-skewed, and the linearity on log-log scales that characterizes theoretical power-law distributions is not evident below a certain threshold[28]. In order to better discern and quantify the power-law features of the distribution, we relied on the statistical framework proposed by Clauset et al. (2009), consisting of few main steps: a) the threshold value and the scaling parameter α of the power-law are estimated; b) a Kolmogorov-Smirnov distance is calculated between the probability distribution of the data and the theoretical distribution defined by the parameters estimated during the first step; such distance is then compared to the distances between probability distributions of synthetic data and the theoretical,

can be written as $\sum_k k^2 p_k - 2\sum_k k\, p_k > 0$, that is as $E(k^2) - 2E(k) > 0$, where $E(k^2)$ and $E(k)$ are respectively the second and first moment of the distribution. The condition is satisfied for power law distributions with scaling parameter $\alpha<3.4788$ (Newman 2003).

[26] A theoretical power law distribution $p_k = ck^{-\alpha}$ is linear on log-log scales. In fact, if we take the logarithm on both sides, we have that $\log(p_k) = \log(c) - \alpha\log(k)$. An alternative way to present data consists in plotting the cumulative distribution function rather than the probability, since the first is more robust against fluctuations that are due to small sample sizes, in the distribution's tail in particular (Clauset et al. 2009).

[27] For issues of computational feasibility, the plots of the distribution of patent connectivity are based on the patent citation network of classes 424 and 514 (the most active classes of our dataset) rather than on the network of all technological classes identified before.

[28] As noticed by Clauset et al. (2009), "in practice, few empirical phenomena obey power laws for all values of x. More often the power law applies only for values greater than some minimum x_{min}. In such cases, we say that the *tail* of the distribution follows a power law" (p. 2).

in order to calculate a "p-value" that, counting the proportion of synthetic distances that are larger than the empirical one, gives a measure of the plausibility of the "power-law hypothesis" for the empirical data at hand (in reality, the details of procedure are more complex; see Clauset et al. (2009))[29]. The test leads to a p-value of 0.65, suggesting that, beyond a threshold, the power law hypothesis cannot be ruled out[30]. The result is consistent with the previously mentioned studies that have suggested that patent indicators such as forward citations obey a power law (Valverde et al. 2006). Moreover, the result draws attention to the studies that have found that the distribution of patent values is also highly skewed (Scherer and Harhoff 2000).

Main results

Descriptive statistics and correlations for independent variables, at the patent level, are shown in Tables 3 and 4 respectively. Table 5 presents the results of the Cox hazard regressions, based on yearly data, in which the failure event is represented by the first instance of patent licensing. Results are presented in terms of hazard ratios. The first specification in Table 5 shows the model with the controls for the alternative sources of uncertainty about the underlying technological innovation and patent issue year fixed effects. The second specification adds the controls for the sources of legal uncertainty. The third specification adds the controls for patent characteristics. The fourth specification adds the remaining controls. The fifth specification, including all controls, adds the scaled connectivity covariate. According to the last specification, higher levels of connectivity, and therefore lower levels of technological uncertainty, are correlated to a more than 11% increase in the hazard rate. The size effect, significant at the 1% level, implies that if the *scaled connectivity* almost triples, this increase is correlated to a more than 11% increase in the hazard rate. As we can notice, the exponentiated coefficient of *post-*

[29] The final step, c), consists in comparing, via likelihood ratio tests, the power-law to alternative distributional hypotheses, such as the exponential or the log-normal.

[30] The test has been performed on the vector of patent connectivity of the 1975-1985 citation network. The fitting procedure has been repeated 100 times, and the value obtained for α is 2.63.

grant is statistically significant but it is negatively correlated to the hazard rate, suggesting that, in our setting, there not seem to be a relevant link between the timing of patent grant and the timing of licensing. In fact, as noticed by Gans et al. (2008), this linkage cannot be generalized, because it may differ significantly across sectors. In a similar way, the exponentiated coefficient of *grant lag* is positively correlated to the hazard rate, suggesting that longer grant lags are correlated to shorter licensing lags. Moreover, we can notice that the exponentiated coefficient of *looked before* is positively correlated to the hazard rate, and the size of the impact is relevant. This seems to suggest that the simple fact that a confidentiality agreement or an option has been signed before may represent an initial signal of quality that, in turn, may boost the hazard rate of licensing. On the other hand, the exponentiated coefficient of the *number of times* a patent has been *looked before* is negatively correlated to the hazard rate. This seems to suggest that, if the patent has been the object of repeated confidentiality/option agreements, not followed by licensing, then maybe an initial signal of quality may have become a signal of technological risk[31], to the detriment of licensing. The exponentiated coefficient of the *number of inventors' patents* is significant but, strangely, is negatively correlated to the hazard rate. The exponentiated coefficient of the *number of technological classes* is also significant, and it is positively correlated to the hazard rate. This seems to indicate that both generality and scope of legal protection may have some positive impact on the hazard rate. Finally, the coefficient of *forward metric* is significant but with the opposite sign. Although puzzling, this results seems to be due to how the metric was conceived[32]. Table 5 presents the results of the Cox hazard regressions with

[31] An option agreement, upon the payment of a fee, gives the licensee firm the right to license the patent within a pre-specified period, and therefore the right to wait until eventual factors of technological risk are resolved.

[32] As mentioned previously, the measure scales the number of forward citations received by the patent until the year that precedes the first license (or, in the case of censoring, the year that precedes the last year of observation) by the average number of forward citations received, until the same year, by patents belonging to the same technological class and to the same cohort. Censored patents, which represent a relevant portion of the database, are observed for longer periods on average: during these periods, they receive more citations than other patents (which may fail quickly due to licensing) and, at the end of these periods, they are not licensed; this may explain why the forward metric is negatively correlated to the hazard rate of licensing.

interaction effects. The fifth specification includes controls, interaction effects and the scaled connectivity covariate. According to the last specification, higher levels of connectivity, and therefore lower levels of technological uncertainty, are correlated to a more than 9% increase in the underlying hazard rate. The size effect, significant at the 10% level, implies that if the *scaled connectivity* almost triples, this increase is correlated to a more than 9% increase in the underlying hazard rate. As we can notice, the *number of claims*, a proxy of the scope of legal protection, becomes significant and it is positively correlated to the hazard rate. Moreover, *backward patent citations*, their *average age* as well as their *minimum age* become significant, with positive, negative and positive correlations to the hazard rate respectively. This suggests that: an increase in the technological maturity of an invention, in terms of more prior art, is positively correlated to licensing; in a similar vein, a decrease in the technological maturity of an invention, in terms of a decrease in the average aging of the prior art[33], is negatively correlated to licensing; finally, a decrease in the age of the oldest prior art is positively correlated to licensing (or, in other words, excessive aging of the oldest prior art is negatively correlated to licensing). Regard to the the interaction effects, the exponentiated coefficients of the interaction of scaled connectivity with the average and the minimum age of backward citations are negative and positive respectively. Moreover, the coefficient of the interaction of the scaled connectivity with the number of non-patent references is negative. The consistency of these results with the hypothesis that technological maturity and closeness to commercial applications may moderate the relationship between technological uncertainty and licensing outcomes, is not very clear. Moreover, the size effects don’t seem to be relevant. In general, we are extremely cautious about our results, because the sample size is small and several of our patent measures are correlated. Additional analyses, not reported here, explored the robustness of the baseline results on a

[33] Note that age variables are in negative values.

smaller sample, in which the patents licensed before their grant date were excluded. The sample consists of 194 patents and of 1499 observations. Both the post-grant and the grant lag regressors have been omitted from these specifications. In the last specification of the entire model without interaction effects, the exponentiated coefficient of scaled connectivity is significant at the 1% level and it is correlated to a 14% increase in the hazard rate[34]. In the last specification of the entire model with interaction effects, the exponentiated coefficient of scaled connectivity is significant at the 1% level, and it is correlated to a 38% increase in the hazard rate. Depending on the availability of larger samples, future analyses should explore the robustness of the baseline results to alternative measures of technological uncertainty based on the "volatility" of scaled connectivity.

6. Discussion

Some preliminary evidence suggests that higher levels of scaled connectivity, and therefore lower levels of technological uncertainty, are correlated to an increase in the hazard of licensing. In this paper, a stylized framework has been proposed in order to make sense of this linkage, shedding light on a possible mechanism that may underlie the relationship between technological uncertainty and markets of technology: in the presence of market incompleteness, technological uncertainty can make problematic the valuation of technological assets. Disagreements on valuation may then lead to less deals and, in general, to a less efficient functioning of markets for technology. The main implication is that, in the presence of fundamental uncertainty, the achievement of market configurations that *tend* to realize the ideal of "market completeness" may improve market functioning. This seems to represent a feasible policy option, as demonstrated by the very last developments in the IP industry, namely the birth

[34] The coefficient of backward citations becomes positive and significant. The coefficients of the number of technological classes, of the number of inventors' patents and of the number of times a patent has been looked before are no longer significant.

of the IPXI, the first centralized financial exchange for patent licenses[35]. Ideally, a market for patents is complete if each patent has a corresponding market and market pricing. The business model of the IPXI is based on the basic intuition that a market, and in general market pricing, cannot form around a patent license until that asset is "commoditised": therefore, the IPXI has split each available patent license in a package of non-exclusive unit licenses, the so-called URL contracts, which are traded on a centralized market that, being open to a significant number of buyers and sellers, should provide adequate levels of thickness, improving the efficiency of the market[36] (McClure 2011). Thanks to its focus on commoditization, the IPXI will likely offer a new "paradigm" for the IP industry (McClure 2011). In fact, according to a traditional business model, the IP market was based on the exchange of a generic patent asset, such as a license, "as a whole" on a bilateral "blind market" (Lemley and Myhrvold 2008). In other words, patents were exchanged as "singularites" (Troy 2012), that is assets characterized by uniqueness, incomparability, and uncommonness (Karpic 2010). As legal constructs, patents are singularities, since they are novel and non-obvious, and therefore unique, by definition (Troy 2012). By splitting each patent license in a package of non-exclusive unit licenses, the URL contracts, the IPXI has "fractioned" the singularity of the legal construct of a patent. However, patents are also singularities as containers of technological knowledge, since technical knowledge is very often

[35] The IPXI has been established by Ocean Tomo, a Chicago-based firm specializing in intellectual property. Its supporting partners include the Chicago Board Options Exchange. On May 4th 2012 the IPXI published a first version of the rulebook that governs how the exchange will function, and it will probably open during the fall of 2013.

[36] Each unit license, available with standardized conditions and at a market based price, gives the owner the right to produce and sell up to a *pre-established* quantity of products protected by the underlying patent. Unit licenses can be acquired, or sold, on an "as-needed" basis, in order to cope with increasing, or decreasing, manufacturing/sales needs (IPXI 2012, website). For example, building on the example provided by Gray (2008), let's imagine that a package of 25 million URL contracts comes to market, where each URL contract gives the right to produce and sell *one* smart-phone incorporating a brand-new multi-touch screen technology patented and owned by a major producer of consumer electronics. Let's imagine that a small producer of smart-phones, still far from the technology frontier of multi-touch devices, decides to incorporate the multi-touch patented technology into its handsets and thus to purchase 2.5 million URL contracts, on the basis of her expectations to produce and sell 2.5 million smart-phones over the next 18 months. The small producer can buy 0.5 million URL contracts now, to cover the production needs over the next three months, and the other 2 million URL contracts during the remaining 15 months. Alternatively, the small buyer can buy 2.5 million URL contracts now, and sell some of them later in case of an early abandonment of the production line. In this light, the URL contract "allows for the creation of a commodity market for trading the units, including the future introduction of derivative products based on those assets" to be used as a hedge against price risk, as well as to acquire risk through speculation and arbitrage strategies (McClure 2011, p. 31).

anchored to a firm-specific context, and it is therefore “costly if not impossible to use elsewhere” (Troy 2012, p. 49), unless the access to complementary technological knowledge is facilitated. This suggests that, in order to foster market participation, the IPXI has to facilitate the access to all the required complementary patent assets[37]. However, as suggested by Gans and Stern (2010), lack of participation to the market may be also due to a lack of complementarity that unfolds over time. In few words, the value of a technology to a firm, and therefore the willingness of the firm to participate to market transactions, may depend on the emergence of complementary technological innovations whose precise form and timing may be serendipitous and difficult, if not impossible, to anticipate (Gans and Stern 2010). Similarly, it may depend on the emergence of complementary technological innovations that may not yet exist (Rosenberg 1996). This suggests that, despite the improvements in terms of efficiency provided by commoditization, technological uncertainty about uses may still play a problematic role for the efficient functioning of markets for technology. If a condition of efficiency is hypothetically realized and, therefore, the price of a patent corresponds to the value as calculated before, then the owner of the patent would simply need to know the *direct* uses of the patent, in order to identify its best value (see p. 10). However, knowledge of uses that are possible for a technology may be undermined by the simple fact that technologies bring with themselves an infinite of uses that cannot be pre-specified and known ex-ante (Bonaccorsi 2011). The mapping between the physical configuration of a technology and the set of all possible functions (uses) lacks inner order and predictability: in fact, given the nature of engineering design problems, “knowledge about physical structures can enumerate exhaustively only all mappings [with functions] that are not feasible, but never those that are” (Bonaccorsi 2011, p. 308). In a similar vein, some recent

[37] The IPXI seems to be aware of the issue. In fact, according to the market rulebook, the IPXI: 1) "solicits comments from market participants and potential licensees...including whether other patents held by third-parties are essential to the unit-base"; 2) "will work with patent pooling organizations to multiple patents for URL Offerings, where necessary and appropriate"; 3) moreover, "URL contracts require that all patented technologies or applications held by Sponsors, at the time of the issuance of later developed, which are essential to the unit-base must be included in the URL contract" (IPXI Market Rulebook).

literature on the evolution of technology has suggested that the uses for which a technology has been optimized for are only a subset of the infinite set of possible causal consequences that it can generate. As underlined by Basalla (1988), inventions offer a range of opportunities, but only few of them will be exploited during their life-time. This suggests that, eventually, technological uncertainty about uses may play a problematic role for the valuation of technological assets and, in turn, for the efficient functioning of markets for technology. At the same time, it may represent a unique opportunity for firms endowed with more absorptive capacity (Cohen and Levinthal 1990). In other words, it may represent an opportunity for firms endowed with the "ability to evaluate" the value of an external technology (Arora and Gambardella 2010b) and to "foresee" new uses, being able to cope, in some extent, with the inherent difficulty to anticipate the serendipitous emergence of technologies and, eventually, to set up profitable strategies that may be offered by arbitrage opportunities. Indeed, as noticed by Barney (1989), when strategic factors markets are characterized by imperfections, firms with better expectations are expected to obtain above normal economic performances.

References

Abbring, J.H., & van den Berg, G.J. 2003. The nonparametric identification of treatment effects in duration models. ***Econometrica***, vol. 71, p. 1491-1518.

Abbring, J.H., & van den Berg, G.J. 2003. A simple procedure for the evaluation of treatment effects on duration models. **Working paper**.

Alcacèr, J., & Gittelman, M. Patent citations as a measure of knowledge flows: the influence of examiner citations. ***The Review of Economics and Statistics***, vol. 88 (4), p. 774-779

Ali, A., & Cockburn, I. 2012. Buyer behavior in markets for technology: technology proximity between firm portfolio and in-licensed patents. **Working paper**.

Arora, A., Fosfuri, A. & Gambardella, A. 2001. **Markets for technology: the economics of innovation and corporate strategy**. MIT Press.

Arora, A., & Gambardella, A. 2010. Ideas for rent: an overview of markets for technology. ***Industrial and Corporate Change***, vol. 19 (3), p. 775-803.

Arora, A., & Gambardella, A. 2010b. The market for technology. In Hall, B., & Rosenberg, N. 2010. **Handbooks of Economics of Innovation**. Elsevier.

Barberà-Tomàs, D., Jimenez-Saez, F., & Castellò-Molina, I. 2011. Mapping the importance of the real world: the validity of connectivity analysis of patent citations networks. ***Research Policy***, vol. 40, p. 473-486.

Barberà-Tomas, D., & Consoli, D. 2012. Whatever works: uncertainty and technological hybrids in medical innovation. ***Technological Forecasting & Social Change***, vol. 79, p. 932-948.

Barney, J.B. 1986. Strategic factor markets: expectations, luck, and business strategy. ***Management Science***, vol. 32 (10), p. 1231-1241.

Batagelj, V. 2003. Efficient algorithms for citation network analysis. **Working Paper**.

Basalla, G. 1988. **The evolution of technology**. Cambridge University Press.

Bellman, R. 1957. **Dynamic programming**. Princeton University Press.

Bonaccorsi, A. 2011. A functional theory of technology and technological change. In Antonelli, C. 2011. **Handbook on the Economic Complexity of Technological Change**. Edward Elgar.

Bresnahan, T.F., & Trajtenberg, M. 1995. General purpose technologies, engines of growth. ***Journal of Econometrics***, vol. 65 (1), p. 83-108.

Clauset, A., Shalizi, C.R., & Newman, M.E.J. 2009. Power law distributions in empirical data. ***SIAM Review***, vol. 51 (4), p. 661-703.

Cohen, W.M., & Levinthal, D.A. 1990. Absorptive capacity: a new perspective on learning and innovation. ***Administrative Science Quarterly***, vol. 35 (1), p. 128-152.

David, P. 1985. Clio and the economics of QWERTY. ***American Economic Review***, vol. 75 (2), p. 332-337.

De Nooy, W., Mrvar, A., & Batagelj, V. 2011. **Exploratory social network analysis with Pajek: revised and expanded second edition**. Cambridge University Press.

Denrell, J., Fang, C., & Winter, S.G. 2003. The economics of strategic opportunity. ***Strategic Management Journal***, vol. 24, p. 977-990.

Dierickx, I., & Cool, K.O. 1989. Asset stock accumulation and sustainability of competitive advantage. ***Management Science***, vol. 35, p. 1504-1511.

Dorfman, R., Samuelson, P.A., & Solow R.M. 1958. **Linear programming and economic analysis**. McGraw Hill.

Dosi, G. 1982. Technological paradigms and technological trajectories. ***Research Policy***, vol. 11 (3), p. 147-162.

Fleming, L, & Sorenson, O. 2004. Science as a map in technological search. ***Strategic Management Journal***, vol. 25, p. 909-928.

Fleming, L. 2007. Breakthroughs and the “long tail” of innovation. ***MIT Sloan Management Review***, vol. 49 (1), p. 69-74.

Fontana, R., Nuvolari, A., & Verspagen, B. 2008. Mapping technological trajectories as patent citation networks. An application to data communication standards. **Working paper**.

Fosfuri, A., & Giarratana, M. 2010. Trading under the buttonwood: a foreword to the markets for technology and ideas. ***Industrial and Corporate Change***, vol. 9 (3), p. 767-773.

Gambardella, A., & Giarratana, M. 2013. General technological capabilities, product market fragmentation, and markets for technology. ***Research Policy***, vol. 42, p. 315-325.

Gans, J.S., David, H.H., & Stern, S. 2008. The impact of uncertain intellectual property rights on the market for ideas: evidence from patent grant delays. ***Management Science***, vol. 54 (5), p. 982-997.

Gans, J.S., & Stern, S. 2010. Is there a market for ideas? ***Industrial and Corporate Change***, vol. 3, p. 805-837.

Gray, C. 2008. A new era in IP licensing: the Unit License Right program. ***The Licensing Journal***, November/December.

Hall, B.H., Jaffe, A.B., & Trajtenberg, M. 2001. The NBER patent citations data file: lessons, insights and methodological tools. **Working Paper**.

Hayek, F.A.V. 1945. The use of knowledge in society. ***American Economic Review***, vol. 35, p. 519-530.

Hegde, D. 2011. Asymmetric information and contracts in the market for ideas: evidence from the licensing of biomedical inventions. **Working paper**.

Herraiz, I., German, D.M., & Hassan, A.E. On the distribution of source code file sizes. **Working paper**.

Jenkins, S.P. 2005. Survival analysis. **Working paper**.

Karpic, L. 2010. **Valuing the unique: the economics of singularities**. Princeton University Press.

Kelly, K. 2011. **What technology wants**. The Penguin Group.

Koopmans, T.C. 1957. **Three essays on the state of economic science**. McGraw Hill.

Lai, R., D'Amour, A., Yu, A., & Fleming, L. 2010. Disambiguation and co-authorships networks of the US Patent Inventor database (1975-2010). **Working paper**.

Lancaster, T. 1990. **The econometric analysis of transition data**. Cambridge University Press.

Lane, D.A. & Maxfield, R.R. 2005. Ontological uncertainty and innovation. ***Journal of Evolutionary Economics***, vol. 15, p. 3-50.

Lanjouw, J.O., & Schankerman, M. 2001. Characteristics of patent litigation: a window on competition. ***Rand Journal of Economics***, vol. 32 (1), p. 129-151.

Lanjouw, J.O., & Schankerman, M. 2004. Patent quality and research productivity: measuring innovation with multiple indicators. ***The Economic Journal***, vol. 114 (495), p. 441-465.

Lemley, M.A., & Shapiro, C. 2005. Probabilistic patents. ***Journal of Economic Perspectives***, vol.19 (2), p. 75-98.

Lemley, M.A., & Myhrvold, N. 2007. How to make a patent market. ***Hofstra Law Review***, vol. 36, p. 257.

Lerner, J. 1994. The importance of patent scope. ***Rand Journal of Economics***, vol. 25 (2), p. 319-333.

Levenhagen, M., Thomas, H., & Porac, J.F. 1990. A model of market formation. **Working paper**.

Mansfield, E. 1995. Academic research underlying industrial innovation: sources, characteristics and financing. ***Review of Economics and Statistics***, vol. 77, p. 55-65.

Martinelli, A. 2010. The dynamics of technological discontinuities: a patent citation network analysis of telecommunication switches. **PhD thesis**.

Martinelli, A. 2012. An emerging paradigm or just another trajectory? Understanding the nature of technological changes using engineering heuristics in the telecommunications switching industry. ***Research Policy***, vol. 41, p. 414-429.

McClure, I. 2011. The value of IP as a commodity. ***Intellectual Asset Management***, May/June.

Mehta, A., Rysman, M., & Simcoe, T. 2010. Identifying the age profile of patent citations: new estimates of knowledge diffusion. ***Journal of Applied Econometrics***, vol. 25, p. 1179-1204.

Merges, R.P., & Nelson, R.R. 1990. On the complex economics of patent scope. ***Columbia Law Review***, vol. 90, p. 839-916.

Myerson, R. 2008. Perspectives on mechanism design in economic theory. ***American Economic Review***, vol. 98, p. 586-603.

Munari, F., & Oriani, R. 2011. **The economic valuation of patents: methods and applications**. Edward Elgar Publishing.

Narin, F., Hamilton, K.S., & Olivastro, D. 1997. The increasing linkage between US technology and public science. ***Research Policy***, vol. 26, p. 317-330.

Newman, M.E.J. 2003. The structure and function of complex networks. **Working paper**.

Oriani, R., & Sobrero, M. 2008. Uncertainty and the market valuation of R&D within a real options logic. ***Strategic Management Journal***, vol. 29, p. 343-361.

Pitkethly, R. 2006. Patent valuation and real options. In Bosworth, D. & Webster, E. 2006. **The Management of intellectual property**. Edward Elgar Publishing.

Razgaitis, R. 2004. US/Canadian licensing in 2003: survey results. ***Journal of the Licensing*** *Executive Society*, vol. 34 (4), p. 139-151.

Rosenberg, N. 1996. Uncertainty and technological change. In Landau, R., Taylor, T. & Wright, G. 1996. **The mosaic of economic growth**. Stanford University Press.

Roth, A.E. 2002. The economist as engineer: game theory, experimentation, and computation as tools for design economics. ***Econometrica***, vol. 70, p. 1341-1378.

Roth, A.E. 2007. The art of designing markets. ***Harvard Business Review***, vol. 85, p. 118-126.

Roth, A.E. 2008. What have we learned from market design? ***Economic Journal***, vol. 118, p. 285-310.

Scherer, F. & Harhoff, D. 2000. Policy implications for a world of skew-distributed outcomes. ***Research Policy***, vol. 29, p. 559-566.

Singh, J., & Fleming, L. 2010. Lone inventors as sources of breakthroughs: myth or reality? ***Management Science***, vol. 56 (1), p. 41-56.

Teece, D.J. 1998. Capturing value from knowledge assets: the new economy, markets for know-how, and intangible assets. ***California Management Review***, vol. 40 (3), p. 55-79.

Temkin, I., & Eldredge, N. 2007. Phylogenetics and material cultural evolution. ***Current Anthropology***, vol. 48 (1), p. 146-153.

Trajtenberg, M. 1990. A penny for your quotes: patent citations and the value of innovations. ***The Rand Journal of Economics***, vol. 21 (1), p. 172-187.

Troy, I. 2012. Patent transactions and markets for patents. **PhD thesis**.

Valverde, S., Solè, R.V., Bedau, M.A., Packard, N. 2006. Topology and evolution of technology innovation networks. ***Physical Review E***, vol. 76, 0561118.

Verspagen, B. 2007. Mapping technological trajectories as patent citation networks: a study on the history of fuel cell research. ***Advances in Complex Systems***, vol. 10 (1), p. 93-115.

Wartburg, I.V., Teichert, T., & Rost, K. 2005. Inventive progress measured by multi-stage patent citation analysis. ***Research Policy***, vol. 34 (10), p. 1591-1607.

Williamson, O.E. 1973. Markets and hierarchies: some elementary considerations. ***American Economic Review***, vol. 63(2), p. 316-325.

Ziedonis, A.A. 2007. Real options in technology licensing. ***Management Science***, vol. 53 (10), p. 1618-1633.

TABLE 1

Set of technological classes of TLO's patents

128, 198, 324, 378, 382, 422, 424, 435, 436, 506, 514, 521, 523, 530,

536, 552, 554,560, 564, 600, 601, 604, 606, 607, 623, 702, 705, 800

FIGURE 1

A "snapshot" of the time evolution of the patent citation network

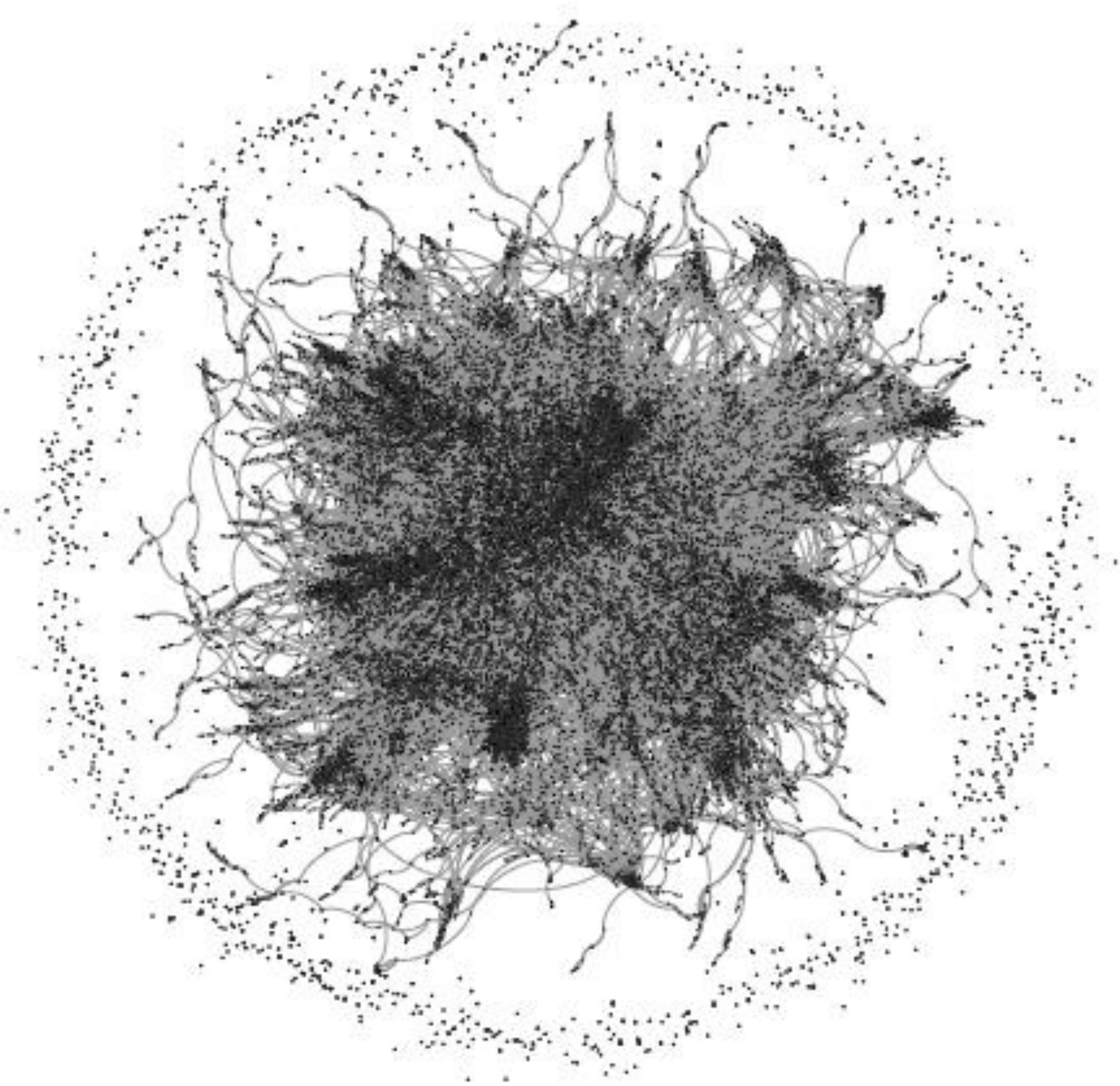


TABLE 2

Time evolution of the patent citation network

Time period	N°Patents	N°Citations	N°Startpoints	N°Endpoints	N°Intermediates	N°Components	Size Largest Component	
1975	139	82	65	74	0	59	6	4.32%
1975-1976	3302	2217	1596	1679	27	1220	28	0.85%
1975-1977	10621	8275	5063	5265	293	2987	242	2.28%
1975-1978	20320	18795	9160	9808	1352	4077	3148	15.49%
1975-1979	27881	28496	12032	13243	2606	4515	10330	37.05%
1975-1980	38238	44062	15670	17585	4983	4709	20882	54.61%
1975-1981	49706	64115	19271	22221	8214	4685	33174	66.74%
1975-1982	59652	83849	22105	25764	11783	4592	43975	73.72%
1975-1983	68872	104122	24516	28697	15659	4421	54077	78.52%
1975-1984	79530	129579	27051	32199	20280	4234	65819	82.76%
1975-1985	90743	158610	29536	35680	25527	4015	78203	86.18%
1975-1986	101905	189136	31926	38978	31001	3800	90256	88.57%
1975-1987	114902	228225	34419	42818	37665	3549	104595	91.03%
1975-1988	127561	268449	36623	46404	44534	3412	117876	92.41%
1975-1989	144014	325986	39492	50926	53596	3215	135145	93.84%
1975-1990	159515	383449	41774	55475	62266	3011	151267	94.83%
1975-1991	176404	451829	44092	60084	72228	2877	168615	95.58%
1975-1992	194417	528818	46463	64811	83143	2741	187119	96.25%
1975-1993	213323	616179	48898	69439	94986	2616	206475	96.79%
1975-1994	231974	717569	51154	73462	107358	2481	225550	97.23%
1975-1995	250868	829395	53373	77551	119944	2362	244837	97.60%
1975-1996	271200	957730	55583	82723	132894	2302	265328	97.83%
1975-1997	293622	1110659	57731	89672	146219	2276	287882	98.05%
1975-1998	321200	1313937	60256	98234	162710	2325	315435	98.21%
1975-1999	349063	1528458	62909	105601	180553	2383	343128	98.30%
1975-2000	376071	1757008	65359	112393	198319	2414	370134	98.42%
1975-2001	405891	2018583	68348	118809	218734	2442	399912	98.53%
1975-2002	436230	2302341	71207	125056	239967	2414	430331	98.65%
1975-2003	466835	2628515	74023	131481	261331	2450	460921	98.73%
1975-2004	492757	2911079	76394	137059	279304	2475	486777	98.79%
1975-2005	514687	3165623	78537	141420	294730	2473	508711	98.84%
1975-2006	541685	3508905	81077	146276	314332	2488	535626	98.88%
1975-2007	566482	3845233	83355	151500	331627	2486	560487	98.94%
1975-2008	589836	4170933	85277	157118	347441	2501	583826	98.98%
1975-2009	615072	4603333	87337	162481	365254	2501	609034	99.02%

FIGURE 2

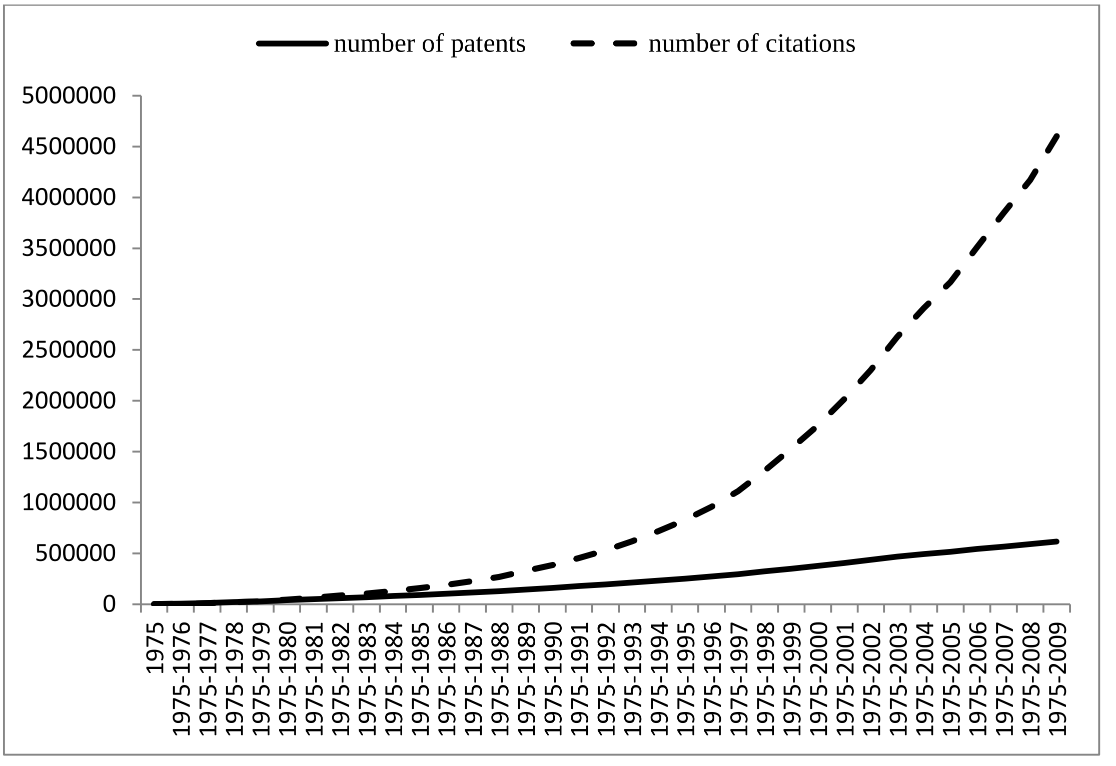


FIGURE 3

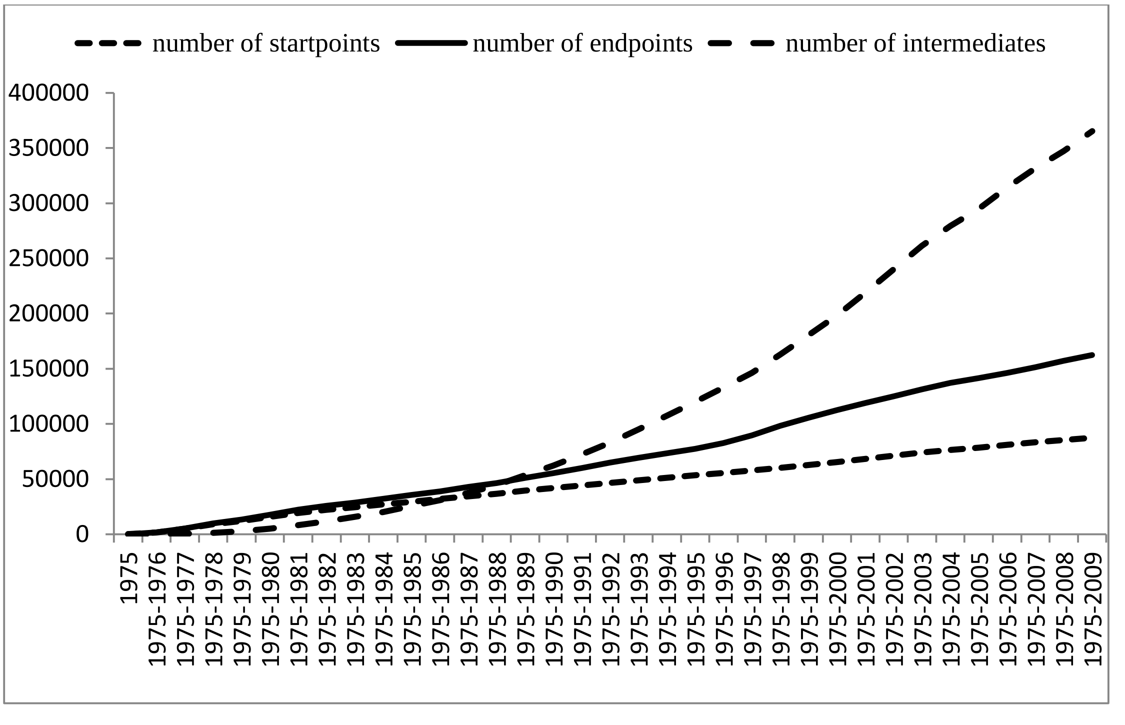


FIGURE 4

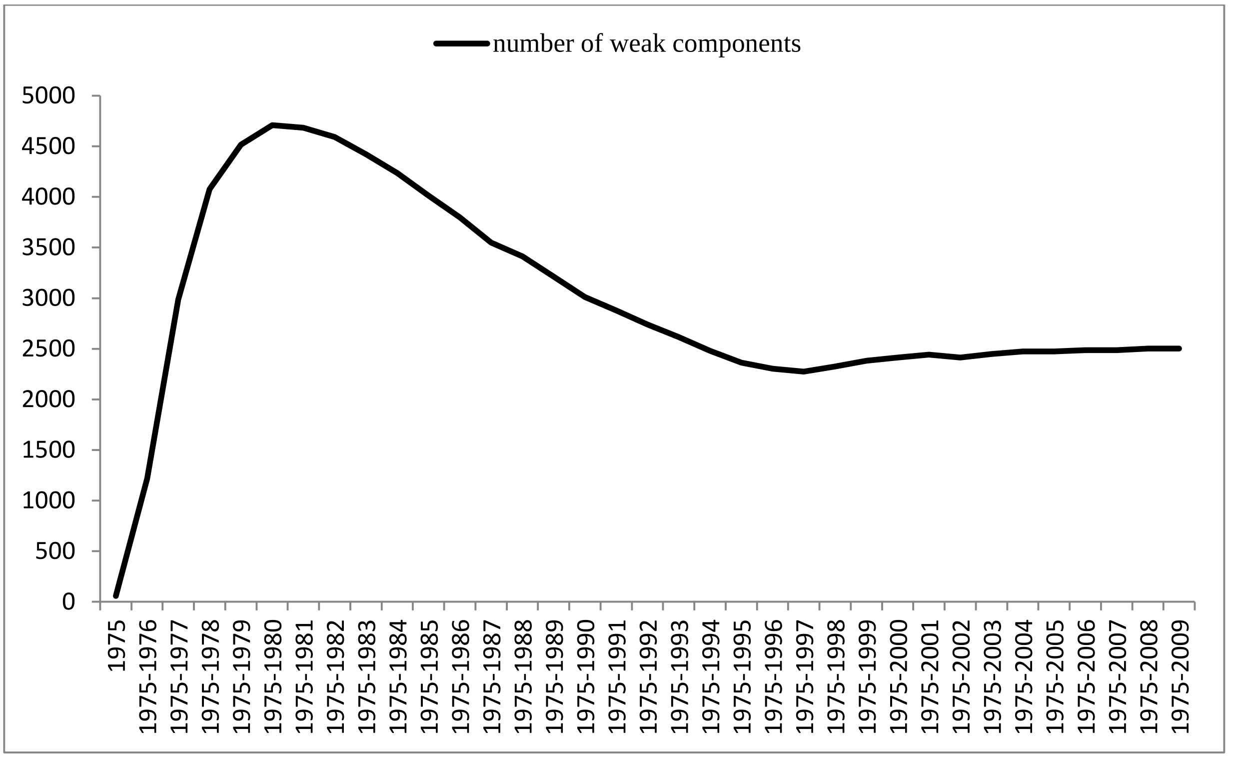


FIGURE 5

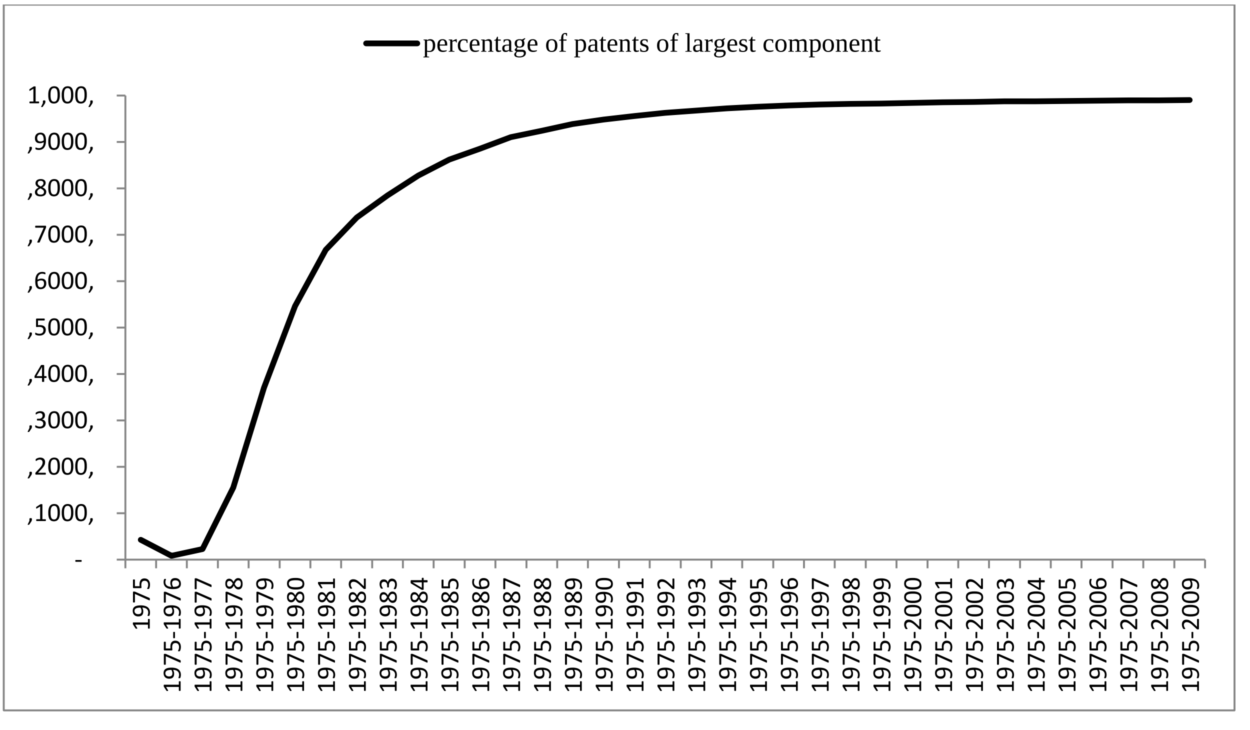
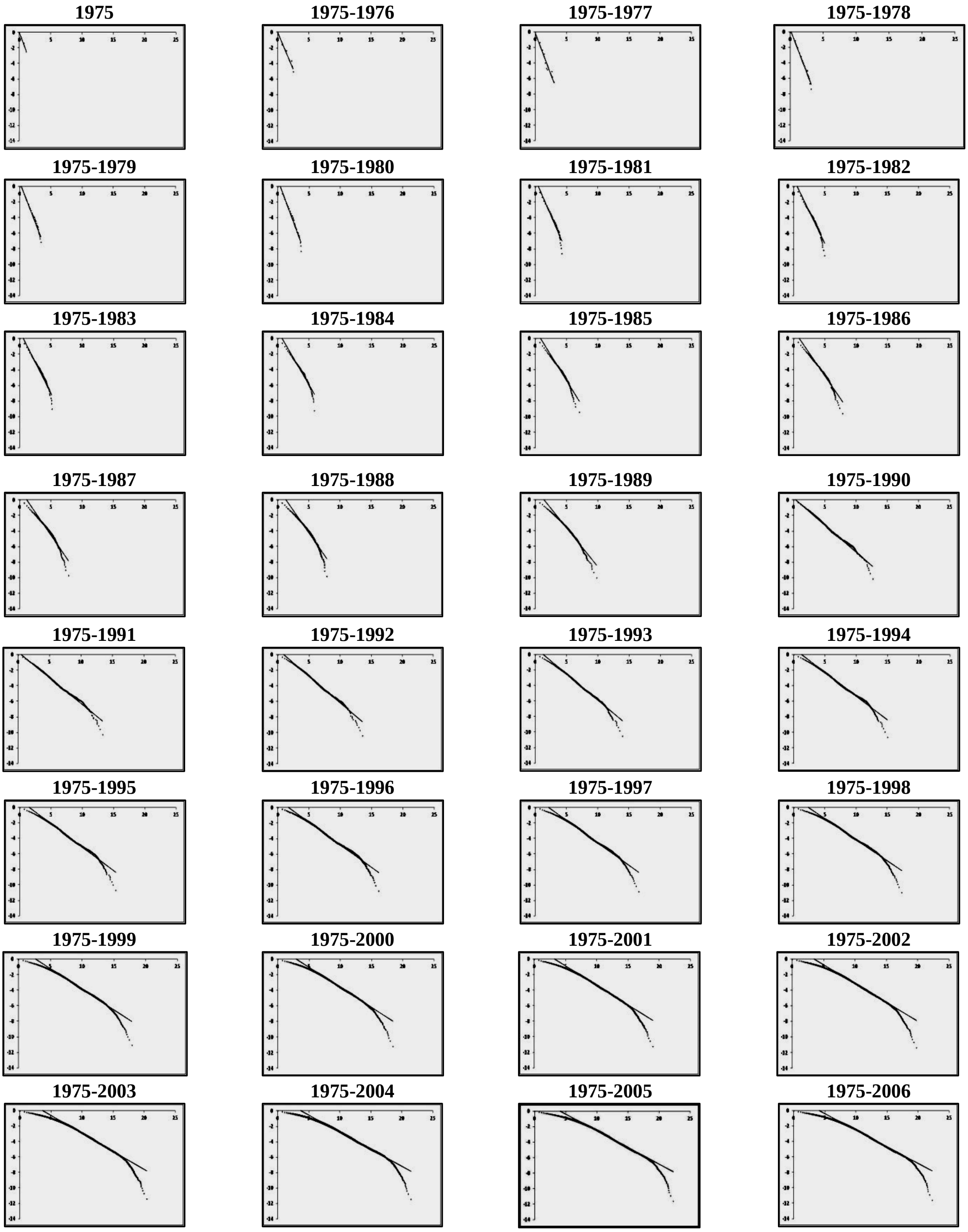
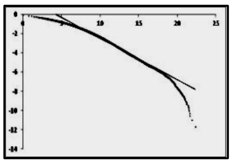


FIGURE 6

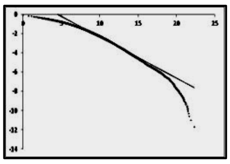
Time evolution of the distribution of connectivity values (log-log scales)



1975-2007



1975-2008



1975-2009

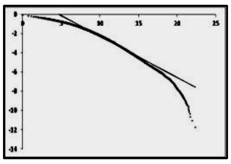


FIGURE 7

Distribution of connectivity for the 1979-2009 period

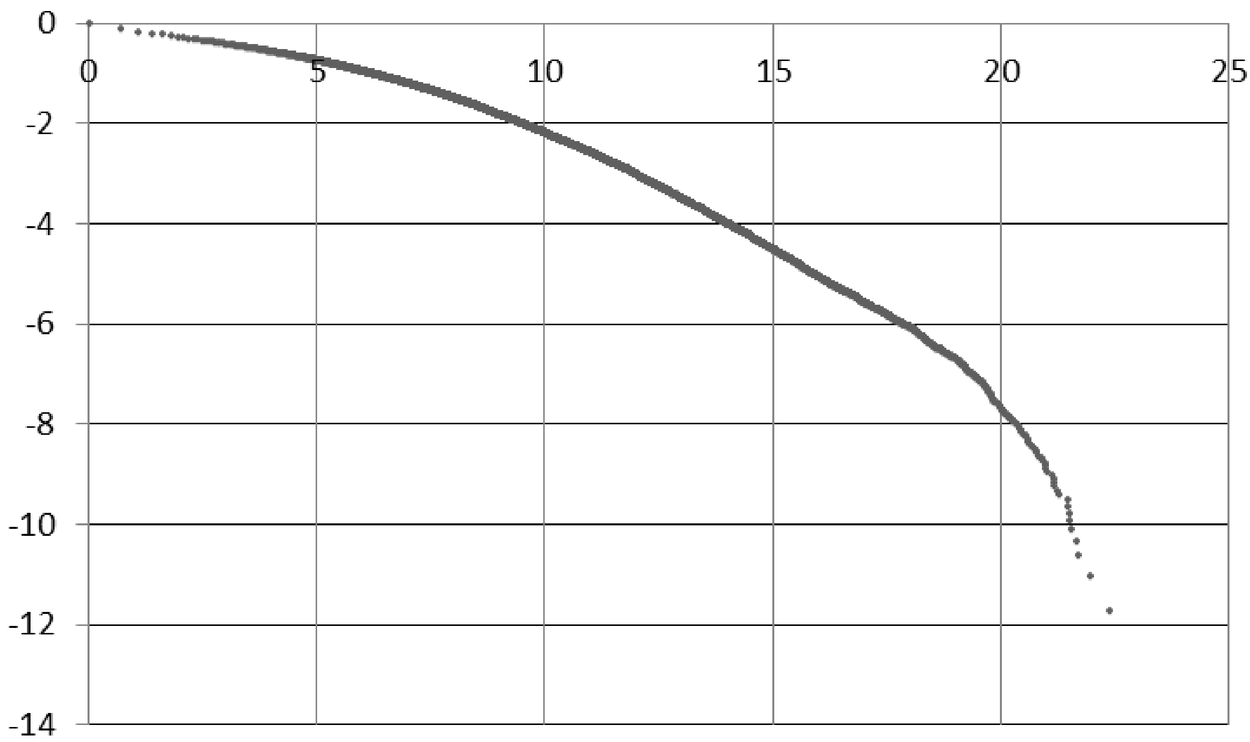


TABLE 3

Descriptive statistics at the patent level

Variable	Mean	Std. Dev.	Min.	Max.
scaled connectivity	-13.480	8.943	-23.025	7.369
backward patent citations	8.766	9.957	0	86
average age of backward patent citations	-8.145	5.741	-32.909	0
minimum age of backward patent citations	-17.654	18.323	-121	0
non-patent references	22.809	25.667	0	146
average age of non-patent references	-8.789	5.682	-41	0
minimum age of non-patent references	-18.906	13.870	-83	0
growth rate of class	0.163	0.243	-0.347	1.732
grant lag	4.849	2.967	1.167	17.487
number of technological classes	1.428	0.716	1	4
number of claims	16.154	12.503	1	93
forward metric	0.401	0.963	0	8.034
number of inventors	2.133	1.120	1	7
number of inventors' patents	14.125	23.817	0	178
looked before	0.104	0.306	0	1
number of times looked before	0.316	1.210	0	11

N=278 patents.

For the same patent, technological uncertainty varies over time. We therefore calculate its average value over time, before calculating the average across patents.

TABLE 4

Correlations at the patent level

Variable+	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1															
2	0.002														
3	0.084	-0.409*													
4	0.017	-0.482*	0.868*												
5	-0.306*	0.087	-0.130	-0.022											
6	0.214*	-0.003	0.085	-0.055	-0.292*										
7	0.208*	-0.071	0.122	0.013	-0.547*	0.759*									
8	-0.092	-0.045	0.011	0.024	0.053	-0.052	-0.054								
9	-0.536*	0.080	-0.225*	-0.036	0.443*	-0.380*	-0.309*	0.112							
10	-0.144	-0.188	0.208*	0.235*	0.065	-0.100	-0.011	0.019	0.001						
11	0.136	0.209*	-0.042	-0.093	-0.045	0.040	-0.006	-0.089	-0.103	-0.063					
12	0.530*	0.038	0.074	0.004	-0.160	0.217*	0.201*	-0.094	-0.192	-0.111	0.143				
13	0.032	0.067	0.046	0.022	0.160	-0.133	-0.135	-0.113	0.013	0.054	0.062	0.219*			
14	0.239*	-0.035	0.022	0.040	-0.018	-0.194	-0.085	-0.086	0.042	-0.019	0.167	0.366*	0.442*		
15	-0.230*	0.237*	-0.247*	-0.204*	0.225*	-0.173	-0.162	0.015	0.120	0.092	-0.051	-0.117	0.169	0.070	
16	-0.208*	0.290*	-0.334*	-0.301*	0.384*	-0.148	-0.264*	0.072	0.208*	0.047	-0.087	-0.097	0.091	-0.014	0.767*

* p<0.01, N=278

+ 1 scaled connectivity; 2 backward patent citations; 3 average age of backward patent citations; 4 minimum age of backward patent citations; 5 non-patent references; 6 average age of non-patent references; 7 minimum age of non-patent references; 8 growth rate of class; 9 grant lag; 10 number of technological classes; 11 number of claims; 12 forward metric; 13 number of inventors; 14 number of inventors' patents; 15 looked before; 16 number of times looked before.

TABLE 5

Cox Proportional Model of the hazard of the first license with shared frailty for technological class; hazard ratios

Variable	1	2	3	4	5
scaled connectivity					1.118***
					(0.036)
backward patent citations	1.025**	1.022**	1.029***	1.027**	1.018
	(0.010)	(0.011)	(0.010)	(0.011)	(0.011)
average age of backward patent citations	0.977	0.995	0.975	0.981	1.002
	(0.033)	(0.036)	(0.033)	(0.035)	(0.036)
minimum age of backward patent citations	1.007	1.003	1.008	1.006	1.002
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
non-patent references	1.004	1.003	0.999	0.999	0.999
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
average age of non-patent references	0.981	0.988	1.002	1.005	0.995
	(0.026)	(0.028)	(0.031)	(0.031)	(0.031)
minimum age of non-patent references	0.999	0.998	1.007	1.005	1.013
	(0.012)	(0.012)	(0.012)	(0.013)	(0.013)
growth rate of class	1.341	1.212	1.536	1.337	1.748
	(0.547)	(0.509)	(0.649)	(0.579)	(0.752)
post - grant		1.356	1.570	1.607	0.138**
		(0.392)	(0.451)	(0.466)	(0.113)
grant lag		1.074*	1.082*	1.112**	1.124***
		(0.042)	(0.044)	(0.047)	(0.048)
number of technological classes			1.276*	1.215	1.266*
			(0.166)	(0.164)	(0.178)
number of claims			1.013*	1.012	1.010
			(0.007)	(0.007)	(0.008)
forward metric			0.0592***	0.0640***	0.0477***
			(0.027)	(0.029)	(0.022)
number of inventors			0.967	0.897	0.931
			(0.097)	(0.093)	(0.097)
number of inventors' patents			0.987*	0.982**	0.981**
			(0.007)	(0.008)	(0.008)
looked before				6.965***	8.087***
				(3.132)	(3.768)
number of times looked before				0.735**	0.742**
				(0.097)	(0.102)

patent issue year fixed effects	included	included	included	included	included
theta	0.4569**	0.4353**	2.11e-16	6.00e-23	0.0361
observations	2448	2448	2448	2448	2448
number of patents	278	278	278	278	278
Wald chi2	32.67	36.57	89.38***	108.14***	115.98***
Log pseudo-likelihood	-737.676	-736.045	-685.747	-676.085	-666.612

*** p<0.01 ** p<0.05 *p<0.1; Gamma shared frailty for technological class; Breslow method for tied failures.

TABLE 6

Cox Proportional Model of the hazard of the first license with interactions effects and shared frailty for technological class; hazard ratios

Variable		**1**	**2**	**3**	**4**	**5**
scaled connectivity	(a)					1.095*
						(0.057)
backward patent citations	(b)	1.042**	1.036**	1.046**	1.052***	1.039**
		(0.016)	(0.016)	(0.018)	(0.019)	(0.018)
average age of backward patent citations	(c)	0.872***	0.853***	0.880***	0.874***	0.907**
		(0.043)	(0.046)	(0.036)	(0.032)	(0.034)
minimum age of backward patent citations	(d)	1.038**	1.040**	1.033**	1.036***	1.028**
		(0.018)	(0.018)	(0.014)	(0.012)	(0.011)
non-patent references	(e)	1.003	1.003	0.996	0.993	0.990
		(0.008)	(0.008)	(0.007)	(0.007)	(0.007)
average age of non-patent references	(f)	0.948	0.945	0.956	0.968	0.977
		(0.033)	(0.033)	(0.034)	(0.027)	(0.026)
minimum age of non-patent references	(g)	1.016	1.009	1.020	1.011	1.018
		(0.019)	(0.019)	(0.023)	(0.018)	(0.018)
growth rate of class	(h)	0.608	0.572	1.074	0.820	0.808
		(0.370)	(0.349)	(1.016)	(0.762)	(0.689)
interaction 1	b*a	1.001	1.001	1.001	1.002**	1.001
		(0.0007)	(0.0007)	(0.0008)	(0.0008)	(0.0007)
interaction 2	c*a	0.993***	0.990***	0.993***	0.992***	0.993***
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
interaction 3	d*a	1.002**	1.002**	1.002***	1.002***	1.002***
		(0.0009)	(0.0009)	(0.0006)	(0.0006)	(0.0006)
interaction 4	e*a	1.000	1.000	1.000	1.000	0.999*
		(0.0004)	(0.0003)	(0.0002)	(0.0002)	(0.0003)
interaction 5	f*a	0.998	0.996*	0.995*	0.996*	0.997
		(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
interaction 6	g*a	1.001	1.001	1.001	1.001	1.001
		(0.001)	(0.001)	(0.001)	(0.0008)	(0.0008)
interaction 7	h*a	0.943**	0.950*	0.977	0.971	0.961
		(0.028)	(0.028)	(0.044)	(0.041)	(0.038)
post - grant			0.350**	0.474	0.410*	0.112*
			(0.165)	(0.257)	(0.214)	(0.145)
grant lag			1.092**	1.102***	1.137***	1.129***
			(0.045)	(0.031)	(0.032)	(0.033)

number of technological classes			1.311***	1.248*	1.259*
			(0.137)	(0.147)	(0.156)
number of claims			1.012**	1.012**	1.010*
			(0.005)	(0.005)	(0.005)
forward metric			0.058***	0.060***	0.044***
			(0.030)	(0.031)	(0.025)
number of inventors			0.960	0.875	0.888
			(0.081)	(0.074)	(0.077)
number of inventors' patents			0.987**	0.980**	0.982**
			(0.006)	(0.008)	(0.008)
looked before				8.552***	9.075***
				(3.963)	(4.720)
number of times looked before				0.763*	0.740*
				(0.123)	(0.125)
patent issue year fixed effects	included	included	included	included	included
theta	0.6185***	0.5118**	+clustered s.e.	+clustered s.e.	+clustered s.e.
observations	2448	2448	2448	2448	2448
number of patents	278	278	278	278	278
Wald chi2	51.89**	61.85***	1.12e+10***	5.59e+10***	34123.21***
Log pseudo-likelihood	-728.378	-722.832	-675.333	-662.680	-659.171

*** p<0.01 ** p<0.05 *p<0.1; Gamma shared frailty or, in case of flat or discontinuous likelihood region, (+) clustered standard errors for technological class; Breslow method for tied failures.