Paper to be presented at the DRUID 2012

on

June 19 to June 21

at

CBS, Copenhagen, Denmark,

BUYER BEHAVIOR IN MARKETS FOR TECHNOLOGY: TECHNOLOGY PROXIMITY BETWEEN FIRM PORTFOLIO AND IN-LICENSED PATENTS

Ayfer Ali
Boston University School of Management
Strategy and Innovation
aali@hbs.edu

Iain Cockburn
Boston University School of Management
Strategy and Innovation
cockburn@bu.edu

Abstract

Markets for technology promise to increase productivity by better allocating innovative capacity across firms. Research on the demand side of these markets, however, has been limited. We use a new dataset of patents available for licensing from a large, innovative academic medical center (AMC) to understand the structure of these markets. Our data includes information on all firms that showed interest in these patents by signing a confidentiality agreement and later decided whether to license or not license the focal technology. Strikingly, we find that of the 285 patents we observe, about 30% of patents are never even looked at, and of those that are looked at about 25% are not eventually licensed.

Because market safety issues are ameliorated in this market, we hypothesize that the lack of demand is due partly to the necessity for complementary technologies in the licensing firm. We measure technology proximity using measures of the overlap of International Patent Classes between the AMC patent and the firm?s own technology. We find that technological proximity is indeed a determinant of the decision to in-license once a patent has been looked at. Firms

prefer to license technologies that are close to their own at the broadest proximity measure level. However, at the more granular level, conditional on broad-level proximity, greater proximity between the licensee?s technology portfolio and the AMC patent makes a license less likely implying that ?close? fit is good but ?very close? fit is detrimental for in-licensing.

Jelcodes:O32,-

BUYER BEHAVIOR IN MARKETS FOR TECHNOLOGY: TECHNOLOGY PROXIMITY BETWEEN FIRM PORTFOLIO AND IN-LICENSED PATENTS

1 Abstract

Markets for technology promise to increase productivity by better allocating innovative capacity across firms. Research on the demand side of these markets, however, has been limited. In this paper, we use a new dataset of patents available for licensing from a large, innovative academic medical center (AMC) to understand the structure of these markets. Our data includes information on all firms that showed interest in these patents by signing a confidentiality agreement and later decided whether to license or not license the focal technology. Strikingly, we find that of the 285 patents we observe, about 30% of patents available for licensing are never even looked at, and of those that are looked at about 25% are not eventually licensed. Firms with a higher number of own patents and older firms are more likely to take a license. A licensed patent is looked at on average 3.24 times, compared to 2.23 times for patents that have been considered but never licensed.

Because market *safety* issues are ameliorated in this market, we hypothesize that the lack of demand is due partly to the necessity for complementary technologies in the licensing firm. We measure technology proximity, as captured by different measures of the overlap of International Patent Classes between the AMC patent and the firm's own technology. We find that technological proximity is indeed a determinant of the decision to in-license once a patent has been looked at but the relationship is more complex. Firms are more likely to license technologies that are close and to their own. While this is true at the broadest proximity measure level, we also note that at the more granular level, conditional on high-level proximity, greater similarity between the licensee's patents and the AMC patent makes execution of a license agreement less likely. This implies that "close" fit is good but "very close" fit is detrimental for in-licensing. Additionally, we offer improved measures for technology proximity between patent portfolios.

2 Introduction

Markets for technologies (MFT), where ideas and early stage technologies are traded, promise substantial allocative efficiencies and opportunities for productivity growth by promoting gains from trade and specialization of innovative labor (Arora and Gambardella, 2010). They are needed when the locus of innovation is outside of the firm best fit to commercialize it. Suppliers of technology can be lone inventors or users uninterested in entrepreneurship, not-for-profit institutions specializing in publicly funded academic research or firms that do not possess the downstream assets to commercialize their technologies in any or all markets (von Hippel, 1976; Bresnahan and Trajtenberg, 1995; Teece, 1986;). On the demand side, potential efficiencies also exist as firms with downstream assets could use their strengths by buying (better) technology from outside instead of (only) relying on their own R&D capabilities (Pisano, 1990).

The potential benefits of markets for technologies can only be realized if they can efficiently provide *stable* matching between each idea for sale and the firm best fit to commercialize it (Gale and Shapley, 1962; Roth, 2008). Market design theorists have pointed out a few characteristics of markets that are needed for such efficiency – *thickness*, *lack of congestion* and *safety*. A market is *thick* if a large proportion of the potential buyers and sellers participate in the market. It is *not congested* if it gives an opportunity to each participant to consider multiple transactions. And, finally, it is *safe* for participants when they choose the market over other ways of transacting and reveal their true preferences without engaging in welfare reducing strategic behavior (Roth, 2008).

Gans and Stern (2010) highlight the three main characteristics of ideas that can prevent markets for technologies from operating efficiently – *idea complementarity, user reproducibility and value rivalry*. *Idea complementarity* is the notion that ideas are only useful in combination with other complementary ideas. Its existence reduces the number of potential matches to any given buyer or seller and increases the requirements for *market thickness*. *User reproducibility* refers to the fact that once disclosed, ideas can easily be reproduced and the buyer can then become a seller or not pay for the idea (Arrow, 1962). *Value rivalry* is the fact that value gained by one user may diminish as others also use the idea. U*ser reproducibility* and *value rivalry* can reduce *market safety* by inducing strategic behavior by the participants which would result in overall reduction of welfare (Roth, 2008)

Strategy research related to Markets for Technologies has concerned itself mostly with market *safety* issues that may force firms to choose to not transact in the market or can make them engage in strategic behavior (Arrow 1962; Pisano, 1990; Gans et al, 2008; Anton and Yao, 1994; Arora and Fosfurri, 2003; Teece, 1986; Zeckhauser, 1995) In this study we are able to abstract from market safety issues and concentrate on *idea complementarity* and its significance for market *thickness.*

In our paper we explore a small market for technologies in the context of technology licensing from an Academic Medical Center (AMC). We observe not only all concluded licenses but also the population of all firms who showed an interest in our sample of patents by signing a confidentiality agreement, evaluative material transfer agreement or an option to an exclusive license and later decided to license or not a patent. This allows us to describe the structure of demand in a market for technology, something that has never been accomplished before.

This market is special in that problems of *safety* and *congestion* in markets for technologies are alleviated or non-existent. Our ideas are patented providing a good degree of appropriability and reducing issues of reproducibility by non-licensees. Second, while our seller is interested in generating income its overarching goal in licensing is to see these technologies commercialized and serving the greater good. As a result, it is willing to negotiate with the buyer and price is not the reason why a license is not concluded with a potential buyer. Licensing officer incentives are aligned with the goal of commercialization, not profit maximizing, and significant resources and effort are expended in attempt to commercialize these inventions. Furthermore, the institution is in the business of research and patient care and will not compete with the licensor downstream. As a result, it has no strategic reasons to withhold invention related information from the potential buyer. Additionally, asymmetric information problems, especially with regard to uncertainty regarding the technology quality are attenuated – the inventions come from one of the largest and most respected research institutions in the world.

Given the elimination of many *market safety* and *congestion* issues however, we are still faced with a puzzle: of our sample of 285, approximately half (47%) are never licensed and some 85 (30%) are never even looked at. Of those that are looked at, but not licensed, the first firm to look arrives, on average, 2.75 years after the patent has been filed, or approximately 4-4.5 years after the invention disclosure. Of those that are licensed at least once, first license occurs at 4 years after patent filing or approximately 5.5 years from invention disclosure on average. A patent that has been looked at, but not licensed, gets 2.23 looks, while one that has been licensed has been considered for licensing by 3.24 firms and licensed by 2.02 on average.

In this study we show that even when market safety issues have been substantially alleviated, markets for technologies remain *thin* in the sense that a large number of inventions remain not only unlicensed but also never looked at. This leads us to focus on the importance of *idea complementarity* for the efficient working of these markets. We explore the topic by asking the following research question: "How does technology complementarity affect firm decision to buy a specific idea in markets for technologies?"

We hypothesize that a firm's decision to license a particular invention is dependent on how technologically close its patent portfolio is to the patent under consideration. However we are also able to test for a more complex relationship between proximity and licensing. Using widely accepted measures of technological distance we show that firms license inventions that are close to what they own at the broad level of measurement indicating that idea complementarity is important in their decision making process. However, we also find that controlling for broad level fit, a very close fit at the more granular level of measurement lowers the likelihood of a license due to potentially duplicating in-house efforts.

3 Literature Review and Hypothesis Development

3.1 Markets for technologies

The volume of trade in markets for technologies has been expanding in recent years. Arora and Gambardella (2010) review recent data from various sources to arrive at a market size of approximately $100 billion globally in 2002 which is about double their earlier estimate of $35-50 billion in the mid-1990s. They also estimate that the market has grown at a higher rate than the average global GDP growth rate in the last two decades (Arora et al. 2001; cf. Athreye and Cantwell, 2007; Robbins, 2006; OECD, 2006). Other survey based studies point to the increasing importance and rate of out and in-licensing by firms (Sheehan et al, 2004; Zuniga and Guellec, 2008; Lichtenthaler and Ernst, 2007; Tsai and Wang, 2009)

There is some evidence, however, that not all technologies supplied get licensed. Using PatVal survey data Gambardella et al. (2007) show that 11% of the firm-owned patents in the sample are licensed but another 7% remain unlicensed even when the firm wants to license them. While there is no information on firm effort in the licensing process, patent quality differences explain the firm's willingness to out-license a particular patent but not whether a license actually occurs. This leads the authors to speculate that it is market and organizational inefficiencies that result in such a licensing shortfall. The result is consistent with other findings that firms are unable to find interested parties in 75% of the cases in which they want to license and are able to conclude licenses for only 4% of the technologies they wish to license. They often cite high search costs for licensees as the reason (Razgaitis, 2004).

3.2 Demand in Markets for Technologies

There is little information regarding firms' demand for outside technologies in the literature. The few available studies are mostly based on survey data on firm practices rather than specific licenses, use different definitions of in-licensing and are difficult to generalize by geography or industry. Using data from a survey on low and medium technology firms from Taiwan, Tsai and Wang (2009) find that 95% of the 753 firms in their sample licensed technology from outside. Rate of in-licensing also appears to differ by country. While attitudes towards in-licensing are similar between Japan and the UK, for example, the incidence of in-licensing is higher in Japan where companies also search more for technology to in-license (Pitkethly, 2001).

Other studies imply passivity on the demand side of these markets and show that the party that initiates the licensing contact is often the supplier (Atuhanegima and Patterson, 1993; Lowe and Crawford, 1984; Ford, 1988). Those who license, however, seem to value the technology that they have acquired. In a survey of firms using university technology, Thursby and Thursby (2004) find that more than half of the respondents use university technology in new product development and 23% note that in-licensed patents from universities were crucial in the development of their products.

Research on the demand side of markets for technologies has focused on the firm's decision to "make" or "buy" outside technologies and the factors that influence that decision. Pisano (1990) shows that the firm's choice of external or internal sourcing of R&D depends on considerations of market *safety*, specifically concerns of appropriability and future hold up due to small-numbers bargaining.

Other studies however show that the success of a strategy of external technology acquisition depends on in-house R&D investment indicating that the two are complements rather than substitutes (Cassiman and Veugelers, 2006; Lowe and Taylor, 1998; Tsai and Wang, 2007). Internal R&D is necessary not only to be able to absorb technologies that the firm has decided to acquire but also to monitor the state of the technology outside the firm's boundaries and evaluate potential technology acquisitions (Rosenberg, 1990; Cohen and Levinthal, 1990; Arora and Gambardella, 1994).

With our study we contribute to this literature by describing the structure of demand in a market for technologies. We use a new dataset of patents from an academic medical center and observe all instances when a firm showed an interest in a technology and its decision to conclude or not a license for that technology later. While the supply side studies have focused on the importance of the product and its attributes to understand this market, our demand-side focused study lets us also explore firm characteristics in the licensing decision. Specifically we are interested in the importance of technology complementarity in firm decision making. We are able to look at complementary technological capabilities in the firm in a very concrete way by observing the patents that the firm already owns and their characteristics. This allows us to answer the question: "Does the technology developed inside the firm influence its decision to acquire a specific outside technology, given interest in the technology."

3.3 The Importance of Complementary Technologies

The importance of complementary assets in firms' technology acquisition decisions has been explored before (Teece, 1986; Pisano, 1990). Two studies by Killing (1978) and Caves et al. (1983) look at how in-licensed technologies relate to a firm's current products and capabilities. They provide descriptive statistics on the type of technologies that firms in-license using a convenience sample of 34 licensee companies in the UK and Canada with over 80 licenses in 1974. They find that 22 percent of the licenses were concluded to strengthen the firm's existing products and 70 percent complemented their current capabilities. However, they only rely on licensee survey reports rather than a technology proximity measure and their definition of proximity relates to the products and firms' capabilities rather than the firms' existing technologies.

Little is known about the influence of a firm's technological portfolio in acquiring innovation from outside. Related studies have looked at the importance of technological proximity for firms’ diversification decisions. Breschi et al. (2003) find that a firm’s diversification decision is path dependent and firms expand into related fields. Building on the resource based view of the firm, Silverman (1999) also shows that firms diversify into areas where their existing technological resources are most relevant. Furthermore, in the context of strategic alliances, firms whose technologies are more similar prior to the alliance partners’ tend to “absorb capabilities” from their partners (Mowery et al., 1996). In fact technological proximity has been used to quantify spillovers (Jaffe, 1987).

In a recent study, Laursen et al. (2010) assume that firms license technologies that are close to what they currently hold and show that firms with a more diverse current portfolio of technology, implying higher “monitoring” and "assimilation” capacity, will license technology that is further away from their current in-house expertise.

Based on findings above that firms may be more willing to diversify into technologically closely related areas, we propose the following hypothesis:

H1: Firms are more likely to license inventions that are close to their own technological portfolio, *ceteris paribus.*

We expect that a firm is better able to know about available technology in an area that is closely related to its current knowledge base, reducing search costs for outside inventions. Furthermore, once such inventions are identified, it will be less costly for the firm to correctly evaluate it and assimilate such outside technology into its current portfolio (Cohen and Levinthal, 1990; Arora and Gambardella, 1994). The firm’s existing technological capabilities will then help it extract the most value from it (Silverman, 1999). In this study we don’t witness a firm’s search for new technology since we only observe firms in the “evaluation” stage. Additional data, in terms of commercialization outcomes will let us observe the process of “value extraction” from the firm’s current resources as well.

More importantly however, technological proximity is necessary because ideas are often only useful with other ideas (Gans and Stern, 2010). Heller and Eisenberg (1998) argue that especially in biomedical research, inventions are so interdependent that when intellectual property rights are held by different entities, commercialization can effectively be blocked in case of coordination failure. Such idea complementarity makes inventions only relevant to a few buyers which further lowers chances of a match in the marketplace. As such, the existence of complementary ideas evidenced by technological proximity will be crucial in a firm's decision to license an invention.

Licensing ideas complementary to the ones that it already owns can greatly benefit a firm that is developing new products. However, we expect that technologies that are very similar to what the firm owns in the sense that they can be substitutes to in-house developed inventions will not be licensed. Let's assume that the quality of the in-house and the in-licensed technology are similar and perfectly observable to the firm. The firm has already incurred significant costs for its version of the invention and expects to receive the full amount of the future revenue stream. If it decided to in-license a very similar technology, however, it would most likely pay future royalties to the licensor. As a result, it would choose not to license.

The difficulty of evaluating such early stage technologies and the costly transfer of tacit knowledge associated with outside inventions will further lower the chances of a firm licensing even if quality of the outside invention was better (Polanyi, 1966; von Hippel, 1994; Agarwal, 2006). Furthermore, it is possible that many firms have incentives that reward company scientists for advancing their own technology to the product stage rather than in-licensed technology. Those same scientists are most likely the ones who are evaluating outside technology as well. Behavioral issues such as the so-called "not-invented-here" syndrome which may cause scientists to evaluate outside inventions as inferior to their own have also been pointed out as potential reasons for preferring in-house technologies (Katz and Allen, 1982). This leads us to our second hypothesis:

H2: Firms are less likely to license inventions that are technologically very close (i.e. potential substitutes) to their own technology portfolio, *ceteris paribus.*

We are able to distinguish between H1 and H2 by using an improved version of a widely accepted measure of technological proximity - the cosine, i.e. the uncentered correlation between the technological classes of a focal patent and the firm patent portfolio (Jaffe, 1986). Instead of USPTO patent classes however we use International Patent Classes that have a nested structure and allow us to measure proximity at different levels of granularity. As suggested by previous scholars, we also improve on proximity measures by using all of the IPC codes assigned to a patent rather than the main IPC code (Benner and Waldfogel, 2008).

4 Data

4.1 Research Setting and AMC data

The main dataset for our study comes from the technology licensing office (TLO) of a large Academic Medical Center (AMC). It contains all 285 AMC patents filed and granted between 1980 and mid-2008 and the associated 307 agreements -- confidentiality agreements, option agreements, licenses or assignments-- signed with interested firms for those patents up to 2011. Patents that are co-owned with other institutions or are the result of sponsored research by companies are not included.

AMC employees are required, by law, to disclose to the TLO any patentable invention developed while at the institution or with funds administered through the institution. The AMC assumes all rights to the intellectual property (IPR) and funds and manages patent prosecution and licensing through the TLO. Income from patents (above costs) is divided in the following way – 25% split between the inventors, 25% to the lead inventor's lab, 25% to the lead inventor's department and 25% to the hospital.

Once a patent is filed, the invention is marketed to potential firms through direct mailing and non-confidential information is made publicly available through the web. Information is also disseminated through inventors' research publications, conference presentations and formal and informal contacts with industry. If a firm decides it is interested in a technology, a confidentiality agreement (CDA) is signed which provides access to in-depth information about the research findings, the IPR protection strategy and patent application.[1] The signing of a CDA does not involve a fee or provide any rights to use the patents. Concurrent CDAs with multiple firms are common.

Some of those firms then return to sign an exclusive option to a license for a certain period of time or a license to the patents. Options include option fees and patent expense reimbursement for the duration and provide no rights to use the patents. Licenses involve a

[1] Note that the American Inventors Protection Act granted the USPTO the right to publish patent applications after 18 months from first filing (priority) date. However, it also gives the right to the applicant to request that the application not be published "but only if the invention has not been and will not be the subject of an application filed in a foreign country that requires publication 18 months after filing (or earlier claimed priority date) or under the Patent Cooperation Treaty" - http://www.uspto.gov/patents/resources/general_info_concerning_patents.jsp - accessed on November 23, 2011

combination of patent expense reimbursements, license fees, maintenance fees, and milestone and royalty payments by the firm. Our data includes all the agreements signed for these patents, categorized into "deals not done" (for CDAs and options) and "deals done" (for licenses). In cases where a firm signs a CDA, then an option and then a license for the same patent, we only select the latest agreement per firm-patent. We end up with 307 agreements (295 useable) and overall 600 patent–agreement pairs (588 useable) because many agreements have multiple patents under them and many patents have been looked at and licensed by multiple firms.

The Licensor is a non-profit institution that does not have the willingness or the ability to compete downstream with potential licensees. It also has a mandate to bring these technologies to market to cure disease and further research rather than just maximize profit. As a result, when licenses are not concluded, it is never because of inability to reach a mutually agreeable price with the company. It is because the potential licensee decided that they were no longer interested in the technology for reasons other than price. This is also seen through qualitative data in the case files - comments by officers about why the potential licensee may not have returned for a license after signing a CDA never list price as the reason. License terms are quite standard based on the technology type but negotiation and variation are possible.

4.2 Company Data

Each company's technology profile at the time of agreement signing was compiled using patent data. For the purposes of this paper, we defined a company's technology position using the IPC patent codes of all the company patents filed before the time of agreement signing (Jaffe, 1987; Silverman, 1999). There are a few studies that show that learning depreciates over time (e.g. Benkard, 2000) and company focus may change and lead to a different technological expertise now from the one many years ago. However, it is not clear how long it takes for such technological expertise to change or expire. Data limitations prevent us from determining which company patents are still valid (rather than expired or abandoned) at time of agreement signing. Future versions of the paper will do robustness checks on whether using only patents filed in the last 5, 10, 15 years, before agreement signing will change our results.

It is important to note that determining patent portfolio in an industry where there are so many mergers and acquisitions (M&A) is difficult. For the top 10 largest pharmaceutical companies, we included the patents of previous top 10 large acquisitions. For example, Hoechst patents were included in the Aventis portfolio after the merger but for smaller firms this was not possible due to data limitations. Some data is available, but incomplete for M&As after 1992 but often M&A deals are described as "acquisition or divestiture of nutrition or vaccine business of firm A" and it is not clear which patents go with such acquisitions or divestitures.

The advantage of using IPC codes rather than USPTO codes is their five nested levels of detail from broad to detailed -- section, class, subclass, main group and subgroup level which lets us measure technology fit at various levels of detail. Furthermore, with the exception of the finest level of measurement – the subgroup level (which we don't use) - patent IPC codes are not laterally nested – i.e. consecutive IPC codes are not subsets of each other except at the subgroup level (excluded).[2]

The main independent variables that we construct using patent IPC codes are our technology proximity measures between the focal AMC patent(s) and the patent portfolio of the interested firm. For this purpose we use the cosine measure pioneered by Jaffe (1986) but with modifications that use all IPC codes on the patent and at different levels of detail - subclass and main group.

The cosine measure calculates the angular distance between two vectors that characterize the firm's and the AMC patent's position in a technology space defined by patent classes. For this purpose we create a technology position vector for a firm's portfolio of patents $F_i=(F_{i1}, F_{i2}, F_{i3}\ldots F_{ik})$, where each 'entry' is the share of a firm's patents' IPC codes in a certain technology class k. A technology position vector, F_j is also created for the specific hospital patent under the firm's agreement. The angular distance between the two vectors is then the measure of technology similarity and it ranges between zero and one, one being a perfect fit and zero being no overlap in technology. It is calculated using the following formula:

$$P_{ij} = \frac{F_i' F_j}{\sqrt{(F_i' F_i)(F_j' F_j)}}, 0 \le P_{ij} \le 1$$

[2] A61F 2/04 is an example of an IPC code. A is the section, 61 is the Class, F is the Subclass, 2 is the Main Group and 04 is the Subgroup. For example A61F 2/04 is a subset of A61F 2/02 at the subgroup level. However, A61F 2 and A61F 3 (at the main group level) are not nested within each other. Please see more at
http://www.wipo.int/ipcpub/#refresh=page&notion=scheme&version=20120101&symbol=A61F0002020000

We modify the main group cosine measure further by computing a "within section" cosine measure constructed based on the above formula, except we exclude all IPC codes in the firm patent portfolio which do not match the AMC focal patent IPC codes at the section levels. This measure looks at proximity of the closest part of a company's technology portfolio to the focal patent so we use it at the nested, main group level. We report these results because we believe they are more appropriate but our proximity measure results have the same sign and significance without the "within section" modifications.

4.3 Control Variables

In keeping with prior literature we use AMC patents' cites to characterize them. Unlike citations in academic articles, citations to previous patents, also referred to as *prior art*, delineate and limit the scope of a patent. If a patent cites a prior patent, it means that it cannot lay a claim to the invention in the previous patent. Patent citations are added to the patent both by the patent filer and the examiner. Normally, the more prior art a patent has, the more incremental it is considered and the more developed the technological area to which it belongs. Conversely, fewer backward citations imply that a patent is pioneering. A count of future cites divided by the number of years since patent grant are used to control for the importance and commercial value of a patent (Trajtenberg, 1990). Measures of patent *originality* and patent *generality,* pioneered by Hall et. al. (2001) are also used as controls. The use the dispersion into different technology groups of cited and citing patents respectively to measure whether the patent is original or general where general patents are those that are a platform for a larger variety of future technologies. We use the number of IPC codes assigned to a patent as a measure of *patent scope* as it has been shown to determine patent valuations and licensing outcomes (Lerner, 1994; Gambardella et. al., 2007; Decheneaux et al., 2008).

Technology age, measured as the time from the AMC patent's priority date[3] to the agreement date, is another control variable. Technology risk is likely to be augmented as the invention gets older because further developments are likely under way. Cohort variables are also included for each five year period since 1980 depending on the priority date of the patent – a patent with a priority date in 1993 belongs to the 1990-1994 cohort.

---

[3] Priority date is the date when the first patent from an invention was filed. That first patent may then be split into different patents or have other material added to it.

An indicator variable about whether the AMC patent protects a medical device was manually coded by reading through the patent claims. The criterion was whether the invention would have required an approval by the FDA as a device in order to be used in the market. Indicator variables for the IPC code section to which the patent belongs were also used as further technology controls. The *lead inventor experience* variable is defined as the number of inventions that the inventor has previously disclosed and patented at this specific technology licensing office (TLO). It is a proxy for the inventor's experience both innovating and navigating the licensing process. It is also a signal to firms of the invention's quality and strong IP rights and hence its commercialization potential.

At the firm level we use count of filed firm patents at time of agreement as a control. Because many of our firms, especially those that appear multiple times in our data undergo mergers and acquisitions, we are not able to include firm fixed effects. Furthermore, even though some firms have multiple agreements, most firms have only one or two.

4.4 Descriptive Statistics

4.4.1 Patent Level Descriptive Statistics

Descriptive statistics for the patent level dataset are included in Tables 1a and 1b. The patents are separated into three different groups – those that were never looked at, those that were looked at but were never licensed and those patents that were licensed at least once. Significance levels of two tailed t-tests of comparisons between the first two groups and the "licensed" group are indicated next to the mean of the variable in the respective group. For example, the stars next to the mean value of the variable "Number of Cited Patents" in the "Never Looked At" group indicate that the difference of the means of the "Number of Cited Patents" between the "Never Looked At" and "Licensed At Least Once" groups is statistically significant.

The "lead inventor experience" variable behaves as expected. Firms can view an inventors' experience as a signal of invention quality and commercialization potential. A larger number of inventions can imply stronger and broader IP rights if the inventor has worked on similar problems before and his previous inventions can be licensed together with the current one. The patent scope variable (Lerner, 1994) is also intended to measure breadth of IP rights and, as expected, it differs significantly between the "never looked at" and "licensed" group as well. As expected, forward cites per year are largest among the licensed patents indicating that those are more important and more inventions build on them.

"Number of Cited Patents," an indicator of how pioneering the technology is, contradicts previous research. Radically innovative patents would have little or no prior art because if a patent cites a previous patent, it means that it builds on it. When we look at the "age of oldest cited patent" statistic, we note that while 92% (140 out of 152) patents in the "Licensed" group cite prior patents, only 77% (66 out of 85) do in the "Never Looked At" group. This indicates that the more pioneering a patent is, the less likely it is to be looked at or licensed and the trend holds over all groups. This also implies that university technology is ahead of industry developments. This implication that firms prefer to license in more established technologies even if that leads to narrower IP rights is confirmed by the average age of patents cited by the AMC patents under the agreements. We note that of those patents that cite at least one prior work those in the "Never Looked At" group cite on average younger patents than those in the other groups and this difference is statistically significant.

Interestingly, while 38% of the patents in the 'never looked at' group are devices, only 29% and 20% in the 'looked at but not licensed' and 'licensed' categories respectively are devices, indicating that devices are much less likely to be licensed.

4.4.2 Patent-Agreement Level Descriptive Statistics

The next level of descriptive statistics is at the agreement – patent level. In this dataset there are 588 observations and each observation is a patent agreement pair. Each agreement is either a 'deal done' or 'deal not done' and some agreements contain multiple patents. Since some patents also have multiple agreements on them, a patent may be in the 'deal not done' column with a certain agreement and in the 'deal done' column with a different agreement.

Our most important independent variables are the proximity measures. The cosine measure at the subclass level defines the overall broad proximity between a firm and a patent while at the main group level it represents proximity at a finer detail. The main group level cosine measure is more appropriate for "within section" fit. Both measures have a range between 0 and 1 with 1 indicating a perfect similarity and 0 indicating no similarity between the focal patent and the firm patent portfolio. We note that our cosine measure at the subclass level is smaller in the "no deal" group than in the "deal" group, indicating that firms license technologies that are closer to the technology that they own. At the cosine main group level measured within the AMC patents' sections, there seems to be no difference between the two groups.

It is important to note that the cosine measures are not defined for agreements with firms that have no patents as they don't have IPC classes for matching with the AMC patent. In Table 2a, we first show the mean of this variable after we replace the cosine measure with a zero indicating no similarity between the patent and firm technology for firms that have no patents. The difference of the means of the cosine subclass level between the "deal" and "no deal" groups is statistically significant in this sample. We then exclude those observations where the cosine is not defined and calculate the means without replacement. The means of the cosines are not significantly different between the groups any more.

In confirmation of our previous result on the age of technology offered for licensing, we find that "deals" are significantly more likely to occur when the technology is older, indicating again that firms prefer more established inventions. They also prefer inventions with higher impact indicated by the difference between means in the two groups along the "citations received" variable. As expected from the patent dataset devices represent a higher percentage of the "deals not done."

5 Results:

5.1 Patent Level Models

Because we have some patents that are "never looked at", some that are "looked at but never licensed" and some that are "licensed at least once", we run a number of analyses to understand what patent characteristics may influence licensing or interest in an invention in general. We run a multitude of regressions using the variables that we have described above. The results are in Tables 4a and 4b.

The first models we run, reported in Table 4a try to explain whether a patent has been looked at by a firm, how many times it has been looked at and how much time has passed before the first look. Each agreement (i.e. a CDA, option to license or a license) is a "look". The second set of models, in Table 4b has licensing outcomes as the dependent variable – whether a patent has been licensed, how many times it has been licensed and time to first license.

Interestingly, we see that the variable that measures the number of citations received, an indication of importance, is not statistically significant in explaining whether a patent is looked at or licensed when other variables are controlled for. However, the number of previous cites is positively related to whether a patent is looked at or licensed, indicating that more established technologies are more likely to be successful in markets for technologies. We also note the importance of the lead inventor experience pointing to the importance of quality signals in a market with a lot of product uncertainty

Similar to the descriptive statistics, we see that patents that are devices are less likely to be successful. This is unexpected given the importance of physicians who would be AMC employees in new device developments (Chatterji et al., 2008). It could, however be due to the fact that new devices do not always require AMC resources or government funding to develop and the best ones may be patented outside of the AMC technology commercialization process. Alternatively, the successful ones could be developed in close collaboration with industry under sponsored research agreements and may thus be excluded from our dataset.

The results are repeated in the next two models – the Poisson and negative binomial- where the dependent variable is the number of times a patent has been considered for licensing or licensed. Because from Table 1a we see that both dependent variables "times looked" at and "times licensed" are over-dispersed with the standard error slightly higher than the mean we conduct a likelihood ratio test which shows that the negative binomial, rather than the Poisson, is the appropriate model.

We also run a zero-inflated version of these models because we have an excess number of zeros in both variables (number of times looked and number of times licensed). This model has two parts –a Poisson model and a logit model. The dependent variable in the Poisson part of the model is the number of times a patent has been looked at or licensed, conditional on being looked at or licensed, respectively. The separate inflation model which is a simple logit explains the excess zeros. Even though the zero inflated negative binomial model would be more appropriate, it doesn't converge and we report results from the zero inflated Poisson. A Vuong test shows that the zero inflated model is more appropriate than the regular Poisson.

The results from the zero-inflated Poisson models are similar to the previous three models. The result that stands out is that patent scope is negative and in some models statistically significant, contradicting previous literature (Lerner, 1994; Decheneaux, 2009; Gambardella et al. 2007).

The last three models are cox hazard models. Here the dependent variable is time to first agreement. Positive results indicate that as the independent variable increases, so does the hazard of an agreement. Note, however, that for ease of interpretation beta coefficients are reported rather than hazard ratios. We report a regular cox hazard model in the first column, then stratified by lead inventor and then by a shared frailty (the equivalent of a random effects model) where each group is identified by a lead inventor and includes said lead inventor's patents. Our results are similar to those from the previous models but are not statistically significant in the same manner.

The models in Table 4b are the same as the ones in Table 4a, except that the dependent variables are related to licensing – i.e. licensed, times licensed, and time to license. The direction and the significance of the results are practically the same as well, except for the scope variable which is no longer negative and statistically significant. Another difference is that the "patent cites per year" variable is now positive and statistically significant in two of our models.

5.2 Patent-Agreement Level Models

Our patent-agreement level analyses are our main results. They test out hypothesis that technological proximity between a focal patent and a potential licensee's patent portfolio is a determinant of whether a license will take place. All our models are logistic regressions (legit) with a dependent variable – "deal" – that is equal to one if a license was signed and zero if a deal was not done. Our full model is the last one in the respective table and it includes all our control variables. We start with a legit of our dependent variable with only the respective fit measures. We then try a bare-bones fit and licensee variable model. For our remaining models we start again with the fit measure and add patent citation based measures such as forward cites, backward cites and scope. We then add patent variables constructed based on cites and IPC codes - originality measure, generality measure. Technology type controls are added next – i.e. A device indicator variable and mutually exclusive dummy variables based on IPC classification by sections. Age variables are included next - technology age at time of agreement and cohort dummies for each 5 year period since 1980 based on the patent priority date. Finally, we add licensee variables -- the number of granted patents that the licensee had filed before the time of the agreement and the square of the number of such patents.

Our main models in Table 5a test our hypotheses above those firms are more likely to license inventions that are similar but not too similar. We operationalize our technology proximity measures using the cosine variables described above. This model includes our entire sample with 588 patent–agreement level observations. We replace the cosine measures with 0s in cases in which they are not defined because the agreement firm has no patents. In table 5b we exclude those observations with undefined cosine measures and are left with 424 observations. As seen from the tables, the sign and statistical significance of our results is unchanged.

We see in these models that a higher technological proximity between the firm and the AMC patent, measured at the IPC code subclass level, is more likely to be associated with a license i.e. a "deal." Holding subclass level proximity constant, however, a higher technological proximity measured at the main group level of the IPC codes is less likely to result in a deal.

The next set of models in Table 6 includes only the first agreement that is signed for a patent whether it is a "deal" or "no deal." We are interested in these results because we are concerned that whether the first agreement is a "deal" or "no deal" may signal patent quality and may influence future licenses, especially non-exclusive licenses of which there may be potentially many per patent. An exclusively licensed invention, on the other hand, takes the patent off the market. Of the 200 patents that have at least one agreement, 6 are excluded because of licensee issues (discussed in the data section) and we are left with 194 patent-first agreement pairs. Our technology proximity results from the previous models still hold in this sample and are statistically significant indicating that the results are robust and are not driven by a few patents that have been licensed multiple times since each patent appears only once in this dataset.

6 Conclusion

In this paper we addressed a gap in the literature on markets for technology by taking a close look at the demand for technology. While this has been attempted in previous papers, our unique dataset that includes not only firms that licensed technologies but also showed interest in them but did not license provides an important control group for our description of the structure of such markets. We showed that proximity matters in the technologies firms decided to license. Our identification comes from variation within a group that showed at least a threshold level of interest in the technology by contacting the licensing office and signing a confidentiality agreement. Future research may expand on this by identifying a larger population of potential buyers in this market based on some other measure of interest – we currently do not include informal channels through which information may have been obtained or inquiries that did not result in signing of a confidentiality agreement.

We also contribute to the literature on measurement of technology proximity by using a new patent statistic – the international patent class which with its nested structure allows for proximity measurement at the broad as well as granular level between different (portfolios of) patents. We also improved on existing measures by including multiple classes rather than just one, resulting in better results (Benner and Weldfogel, 2008). Further comparison and validation of these new measures is in order.

Ultimately the real question is whether these technologies make it to the product market once they are licensed and how the technology proximity, either at the broad or the granular level influences that outcome. It would be interesting to know whether in-licensed technologies that are very close to the licensee's in-house developed technology are strategically shelved or perhaps not absorbed by the firm due to behavioral resistance to outside innovations, the so called "not-invented-here" syndrome. (Katz and Allen, 1982; Thursby and Thursby, 2004) We view this paper as a first step in that direction.

Bibliography (incomplete):

Agarwal, AK (2006), "Engaging the Inventor: Exploring Licensing Strategies for University Inventions and the Role of Latent Knowledge" *Strategic Management Journal*, 27 (1), 63-79

Anton, J.J. and D.A. Yao (1994), "Expropriation and Inventions: Appropriable Rents in the Absence of Property Rights." American Economic Review, vol. 84 190-209.

Arora, A., Fosfuri, A., and Gambardella, A. (2001) Markets for Technology – The Economics of Innovation and Corporate Strategy, MIT Press, Cambridge, Mass.

Arora, A. and A. Gambardella (1994), 'Evaluating technological information and utilizing it: scientific knowledge, technological capability, and external linkages in biotechnology,' *Journal of Economic Behavior & Organization*, 24(1), 91–114.

Arora, A. and A. Gambardella (2010), "Ideas for Rent: An Overview of Markets for Technology," *Industrial and Corporate Change*, vol. 19 (3) 775-803

Arora, A. and A. Fosfuri (2003), 'Licensing the market for technology,' *Journal of Economic Behavior and Organization*, 52, 277–295.

Arrow, K. (1962), "Economic Welfare and the Allocation of Resources for Invention," in: *The Rate and Direction of Inventive Activity: Economic and Social Factors*, National Bureau of Economic Research, Inc., 609-626

Athreye, S. and J. Cantwell (2007), 'Creating competition? Globalisation and the emergence of new technology producers,' *Research Policy*, 36(2), 209–226.

Atuahenegima, K (1993), "Buying Technology for Product Development in Smaller Firms", *Industrial Marketing Management*, 22 (3), 223-232

Atuahenegima, K. (1992), Inward Technology Licensing As an Alternative to Internal Research-and-Development in New Product Development - a Conceptual-Framework, *Journal of Product Innovation Management*, 9(2), 156-167

Atuahenegima, K. and P. Patterson (1993), "Managerial Perceptions of Technology Licensing as an Alternative to Internal Research-and-Development in New Product Development - an Empirical-Investigation," *R & D Management*, 23 (4), 327-336

Benkard, L. (2000), "Learning and Forgetting: The Dynamics of Aircraft Production", *American Economic Review,* vol. 90 (4), 1034 – 1054

Benner, M. and J. Waldfogel (2008), "Close to you? Bias and precision in patent-based measures of technological proximity", *Research Policy*, vol.37, 1556-1567

Breschi S., F. Lissoni F and F. Malerba (2003), "Knowledge-relatedness in firm technological diversification," *Research Policy*, 32 (1), 69-87

Bresnahan, T. and M. Trajtenberg (1995), "General purpose technologies 'Engines of growth'?" *Journal of Econometrics*, vol. 65 (1) 83-108

Cassiman B. and R. Vergeulers (2006), In search of complementarity in innovation strategy: Internal R&D and external knowledge acquisition, *Management Science*, 52 (1), 68-82

Caves, R., H. Crookell and JP Killing (1983), "The Imperfect Market for Technology Licenses," *Oxford Bulletin Of Economics And Statistics*, vol. 45 (3)

Cockburn, I.M., M.J. MacGarvie and E. Muller (2010) Patent thickets, licensing and innovative performance, *Industrial And Corporate Change*, 19 (3), 899-925

Cohen, W. and D. Levinthal (1990), "Absorptive Capacity: A New Perspective On Learning and Innovation". *Administrative Science Quarterly*, 35(1), 128-152.

Dechenaux E., B. Goldfarb, S. Shane and M. Thursby (2008), Appropriability and Commercialization: Evidence From MIT Inventions , *Management Science*, 54 (5), 893-906

Ford, 1988

Gale, D. and L. S. Shapley (1962), "College Admissions and the Stability of Marriage", *American Mathematical Monthly* 69, 9-14

Gambardella, A., P. Giuri and A. Luzzi (2007), "The market for patents in Europe," Research Policy, 36(8), 1163–1183.

Gans, J., D. H. Hsu and S. Stern (2008), "The impact of uncertain intellectual property rights on the market for ideas: evidence from patent grant delays," Management Science, 54(5), 982–997.

Gans, J. and S. Stern (2010), "Is There a Market for Ideas", *Industrial and Corporate Change*, vol. 19 (3), 805-837

Hall, B., A. Jaffe and M. Trajtenberg (2001), "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools," NBER Working Paper No. 8498

Heller, MA and RS Eisenberg (1998), Can Patents Deter Innovation? The Anticommons in Biomedical Research, *Science, v*ol. 280 (5364), 698-701

Henderson, R., A. Jaffe & M. Trajtenberg, (1998) "*Universities* as a Source of Commercial *Technology*: A Detailed Analysis of *University* Patenting, 1965-1988", *Review of Economics and Statistics,* Vol. 80(1), 119-127.

Jaffe, Adam B. (1986) "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits and Market Value," American Economic Review, 76, 984–1001.

Katz, R. and T. Allen (1982), Investigating the Not Invented Here (NIH) syndrome: A look at the performance, tenure, and communication patterns of 50 R & D Project Groups, *R&D Management*, 12 (1), 7-20.

Killing, JP (1978), "Diversification through Licensing", R&D Management, 8(3), 159 – 163

Kurokawa S. (1997), "Make-or-buy decisions in R&D: Small technology based firms in the United States and Japan," *IEEE Transactions on Engineering Management*, 44 (2), 124-134

Laursen, K; M, Leone, and S. Torrisi, (2010) "Technological exploration through licensing: new insights from the licensee's point of view," *Industrial and Corporate Change*, vol. 19 (3) 871 - 897

Lerner, J (1994), The Importance Of Patent Scope - An Empirical-Analysis , *RAND Journal Of Economics*, 25 (2), 319-333

Lichtenthaler, U. and H. Ernst (2007), 'External technology commercialization in large firms: results of a quantitative benchmarking study,' R&D Management, 37(5), 383–397.

Lowe and Crawford, 1984

Lowe J. and P. Taylor (1998) R & D and Technology Purchase through License Agreements: Complementary Strategies and Complementary Assets, *R & D Management*, 28 (4), 263-278

Mowery D., J. Oxley and B. Silverman (1996),Strategic alliances and interfirm knowledge transfer, *Strategic Management Journal*, 17 (SI), 77-91

Nerkar A. and S. Shane (2007), Determinants of invention commercialization: An empirical examination of academically sourced inventions, *Strategic Management Journal*, vol. 28, 1155-1166

OECD

Patel, P and K. Pavitt (1994) The continuing, widespread (and neglected) importance of improvements in mechanical technologies, *Research Policy,* 23, 533-545

Pisano, G. (1990) "The R&D Boundaries of the Firm: An Empirical Analysis," *Administrative Science Quarterly*, Vol.35 (1) 153-176

Pitkethly, R.H. (2001), "Intellectual property strategy in Japanese and UK companies: patent licensing decisions and learning opportunities", *Research Policy*, v.30, 425–442

Polanyi, M (1966), The Tacit Dimension, Garden City, N.Y.: Doubleday, 1966.

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

Robbins, C. (2006), 'Measuring payments for the supply and use of intellectual property,' Bureau of Economic Analysis, U.S. Department of Commerce: Washington, DC.

Rosenberg, N. (1990). "Why Do Firms Do Basic Research (With Their Own Money)?" *Research Policy, 19*(2): 165-174.

Roth, A. E. (2008) "What have we learned from market design?" Hahn Lecture, *Economic Journal*, vol. 118, 285–310

Sheehan, J., C. Martinez and D. Guellec (2004), 'Understanding business patenting and licensing: results of a survey,' OECD: Paris.

Silverman B. (1999), "Technological resources and the direction of corporate diversification: Toward an integration of the resource-based view and transaction cost economics ," *Management Science*, 46 (8), 1109-1124

Teece, D. (1986), “Profiting from Technological Innovation: Implications for integration, collaboration, licensing and public policy”, *Research Policy*, vol. 15, 285-305

Thursby J.G. and M.C. Thursby (2004), “Are faculty critical? Their role in university-industry licensing”, *Contemporary Economic Policy*, 22 (2), 162-178

Trajtenberg, M. (1990) A Penny for your Quotes - Patent Citations and the Value Of Innovations, *RAND Journal of Economics*, 21 (1), 172-187

Tsai, K. and J.Wang, (2007), Inward technology licensing and firm performance: a longitudinal study, *R & D Management*, 37 (2), 151-160

Tsai, K. and J.Wang, (2009) “External technology sourcing and innovation performance in LMT sectors: An analysis based on the Taiwanese Technological Innovation Survey”, *Research Policy* vol. 38, 518–526

von Hippel, E. (1976) "The Dominant Role of Users in the Scientific Instrument Innovation Process," *Research Policy*, vol. 5, no. 3 (July):212-39.

von Hippel, E. (2004), Democratizing Innovation (Cambridge: MIT Press)

Zeckhauser R. (1996), “The Challenge of Contracting for Technological Information," Proceedings of the National Academy of Sciences,” vol. 93, 12743-12748

Zuniga, M. P. and D. Guellec (2008), ‘Survey on patent licensing: initial results from Europe and Japan,’ OECD: Paris.

1. Patent Level Descriptive Statistics

Table 1a: This table contains descriptive statistics for the AMC patents that are to be licensed. Two sided t-tests of difference in means between "Never Looked At" and "Licensed at Least Once"; significance indicated in the "Never Looked At" mean column; Two sided t-test of difference in means between "Looked At, Not Licensed" and "Licensed At Least Once"; significance indicated in the "Looked At, Not Licensed" mean column; *** p<0.01, ** p<0.05, * p<0.1; Variable descriptions in Appendix A.

	Never Looked At					Looked At, Not Licensed					Licensed At Least Once				
	N	Mean	Std Dev	Min	Max	N	Mean	Std Dev	Min	Max	N	Mean	Std Dev	Min	Max
Lead Inventor Experience	**85**	4.61***	3.43	1	15	**48**	6.02	4.61	1	17	**152**	7.19	4.77	1	25
Patent Scope	**85**	2.43***	1.38	1	9	**48**	2.87	1.48	1	6	**152**	3.71	3.49	1	19
Cites Per Year	**85**	0.78	1.18	0	6	**48**	0.33**	0.59	0	3	**152**	0.96	1.83	0	19
Cites First Two Years	**85**	0.42	0.90	0	6	**48**	0.33	0.88	0	4	**152**	0.47	0.85	0	5
Originality	**85**	0.64	0.32	0	1	**48**	0.67	0.29	0	1	**152**	0.68	0.23	0	1
Generality	**62**	0.62	0.20	0	1	**25**	0.72***	0.22	0	1	**126**	0.57	0.24	0	1
Number of Cited Patents	**85**	6.57***	6.97	0	29	**48**	6.83**	7.27	0	37	**152**	10.16	10.19	0	54
Share of Agreements that are Licenses											**152**	0.79	0.28	0.125	1
Times Looked						**48**	2.22**	1.85	1	8	**152**	3.24	3.12	1	15
Times Licensed											**152**	2.02	1.66	1	11
Time to First Agreement						**48**	2.75	2.45	0.00	7.98	**152**	3.44	3.68	0.00	12.76
Time to First License											**152**	4.20	3.84	0.00	18.48
Age of Oldest Cited Patent	**66**	15.69**	11.48	2.34	55.56	**45**	17.29	12.66	1.36	57.55	**140**	21.55	17.34	1.09	74.81
Mean Age of Cited Patents	**66**	7.53***	3.95	1.94	18.73	**45**	8.71	4.31	1.36	16.99	**140**	9.94	4.84	1.09	25.21
Median Age of Cited Patents	**66**	6.38***	3.52	1.74	14.46	**45**	7.74	3.92	1.36	16.99	**140**	8.68	4.25	1.09	22.84

Table 1b: This table contains descriptive statistics for the AMC patents that are to be licensed. Two sided t-tests of difference in means between "Never Looked At" and "Licensed at Least Once"; significance indicated in the "Never Looked At" column; Two sided t-test of difference in means between "Looked At, Not Licensed" and "Licensed At Least Once"; significance indicated in the "Looked At, Not Licensed" column; *** p<0.01, ** p<0.05, * p<0.1; Variable descriptions in Appendix A.

Group	Never Looked at (N=85)	Looked At, Not Licensed (N=48)	Licensed (N=152)
	Proportion of "Never Looked At" Patents in Group	Proportion "Looked At, Not Licensed" Patents in Group	Proportion of "Licensed" Patents in Group
Patent Priority Date in 1981-85	0.07	0.00*	0.07
Patent Priority Date in 1986-90	0.28	0.02***	0.25
Patent Priority Date in 1991-95	0.23***	0.08***	0.43
Patent Priority Date in 1996-00	0.30*	0.66***	0.20
Patent Priority Date in 2001-04	0.10*	0.22***	0.05
Device	0.38***	0.29	0.20

2. Patent Agreement Level Descriptive Statistics

Table 2a: This table contains descriptive statistics for the agreement-patent level data. Each observation corresponds to a patent-agreement pair - a patent can have multiple agreements and each agreement can be associated with multiple patents. Patent level measures correspond to the hospital patent which is under the agreementTwo sided t-tests of difference in means between "Deal " and "No Deal" ; significance indicated in the "No Deal" mean column; *** p<0.01, ** p<0.05, * p<0.1; Variable descriptions in Appendix A.

	No Deal					Deal				
Variable	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Cosine Subclass Level†	287	0.24**	0.31	0	1	301	0.30	0.31	0	1
Cosine Subclass Level (no replacement)††	198	0.35	0.32	0	1	226	0.39	0.30	0	1
Within Section Cosine, Group Level†	287	0.22	0.31	0	1	301	0.22	0.28	0	1
Within Section Cosine, Group Level (no replacement)††	198	0.33	0.32	0	1	226	0.29	0.28	0	1
Patent Scope	287	3.06*	1.75	1	10	301	3.45	2.94	1	19
Patent Citations Received Per Year	287	0.57***	0.74	0	3.67	301	0.95	1.79	0	18.91
Patent Citations First Two Years	287	0.38	0.74	0	5	301	0.45	0.87	0	5
Number of Patents Cited	287	10.88	10.66	0	54	301	10.53	10.22	0	54
Patent Originality	287	0.71	0.22	0	1	301	0.71	0.20	0	1
Patent Generality	194	0.60	0.26	0	1	246	0.58	0.24	0	1
Device	287	0.15	0.36	0	1	301	0.12	0.33	0	1
Technology Age in Years	287	5.02***	4.38	0	18.11	301	6.64	4.75	0	19.35
Lead Inventor Experience	287	7.85	5.10	1	22	301	7.27	4.52	1	25
Firm R&D age	198	17.04***	11.47	0.06	43.27	226	21.20	12.01	0.08	46.17
Number of Firm Patents	287	985.01***	2657.38	0	16392	301	2204.26	4227.66	0	15266
Numbers of AMC Patents Under Agreement	287	2.75***	1.70	1	6	301	4.20	4.36	1	19

†Observations for which a cosine measure was not defined because the firm has no patents of its own were assigned a cosine measure of 0.

††Observations for which a cosine measure was not defined were excluded from this calculation.

Table 2b: This table contains descriptive statistics for the agreement-patent level data. Each observation corresponds to a patent-agreement pair - a patent can have multiple agreements and each agreement can be associated with multiple patents. Patent level measures correspond to the hospital patent which is under the agreement. Two sided t-tests of difference in means between "Deal " and "No Deal" ; significance indicated in the "No Deal" mean column; *** p<0.01, ** p<0.05, * p<0.1; Variable descriptions in Appendix A.

	No Deal (N=287)	Deal (N=301)
Group	Proportion of "No Deal" Patent Observations in Group	Proportion of "Deal" Patent Observations in Group
Firm Has No Patents	0.31*	0.25
Firm Has >0 and < 500 Patents	0.47	0.43
Firm Has >500 Patents	0.22***	0.32
Patent Priority Date in 1981-85	0.13	0.13
Patent Priority Date in 1986-90	0.13***	0.25
Patent Priority Date in 1991-95	0.09***	0.36
Patent Priority Date in 1996-00	0.53***	0.25
Patent Priority Date in 2001-04	0.13***	0.02

4. Patent Level Models

Table 4a. This table contains models of patent consideration for licensing i.e. "looks at patent" based on patent characteristics. P-values reported under coefficients.

	LOGIT	**Poisson**	**Negative Binomial Regression**	**Zero Inflated Poisson**		**Cox Hazard Model**	**Cox Hazard Model** Stratified by Lead Inventor	**Cox Hazard Model** Frailty by Leadinventor
VARIABLES	Looked (looked=1)	Times looked	Times looked	Times looked	Inflate (Times looked=0)	Time to Look	Time to Look	Time to Look
Number of Patents Cited	0.0837***	0.0311***	0.0346***	0.0225***	-0.110*	0.013	0.020	0.0219*
	0.001	0.000	0.000	0.000	0.096	0.113	0.276	0.083
Patent Citations Received Per Year	0.023	0.042	0.058	0.039	-0.737	0.006	0.342*	-0.009
	0.818	0.134	0.230	0.162	0.292	0.875	0.0621	0.903
Patent Scope	0.103	-0.0388**	-0.037	-0.0456**	-0.16	-0.0952***	-0.007	-0.031
	0.263	0.045	0.234	0.019	0.405	0.005	0.919	0.515
Lead Inventor Experience	0.118***	0.0349***	0.0380**	0.011	-0.293***	0.000	0.115	0.030
	0.002	0.000	0.011	0.265	0.00264	0.978	0.144	0.348
Patent Originality	0.318	0.208	0.309	0.560***	2.399	0.0381	-0.976	-0.107
	0.582	0.261	0.298	0.009	0.186	0.909	0.189	0.833
Device	-1.047**	-0.824***	-0.844***	-0.972***	-12.59	-0.125	-34.3	-0.523
	0.018	0.000	0.000	0.000	0.977	0.58	1.000	0.215
Patent Priority Date in 1981-85	-0.901	0.646***	0.305	0.749***	3.522	0.346	-1.815	0.614
	0.245	0.001	0.394	0.001	0.111	0.414	0.285	0.378
Patent Priority Date in 1986-90	-0.481	0.062	-0.132	-0.006	1.798	-0.286	-3.095**	-1.127**
	0.389	0.743	0.633	0.976	0.376	0.361	0.041	0.029
Patent Priority Date in 1991-95	0.163	-0.085	-0.139	-0.165	1.193	-0.711**	-2.789**	-1.177**
	0.765	0.642	0.601	0.390	0.565	0.023	0.050	0.021
Patent Priority Date in 1996-00	-0.001	0.413**	0.415	0.461**	2.454	0.307	0.105	0.177
	0.998	0.016	0.109	0.010	0.216	0.298	0.924	0.708
IPC class dummies	yes	yes	yes	yes	yes	yes	yes	yes
Constant	-0.923	-0.025	-0.132	0.109	-2.673			
	0.380	0.939	0.785	0.757	0.323			
lnalpha			-0.457***					
			0.006					
Observations	285	285	285	285	285	200	200	200
Groups								86

*** p<0.01, ** p<0.05, * p<0.1

Table 4b. This table contains models of licensing based on patent level dependent characteristics. P-values under coefficients.

	LOGIT	**Poisson**	**Negative Binomial Regression**	**Zero Inflated Poisson**		**Cox Hazard Model**	**Cox Hazard Model** Stratified by Lead Inventor	**Cox Hazard Model** Frailty by Lead Inventor
VARIABLES	Licensed (licensed=1)	Times licensed	Times licensed	Times licensed	Inflate (Times licensed=0)	Time to License	Time to License	Time to License
Number of Patents Cited	0.0926***	0.0307***	0.0305***	0.008	-0.286***	0.010	0.036	0.006
	0.000	0.000	0.000	0.224	0.001	0.280	0.140	0.694
Patent Citations Received Per Year	0.048	0.0695**	0.0819*	0.038	-2.630**	-0.027	0.339	-0.060
	0.650	0.031	0.066	0.288	0.012	0.567	0.125	0.394
Patent Scope	0.143	0.005	0.007	-0.020	-0.397	-0.021	0.031	0.029
	0.111	0.827	0.823	0.435	0.164	0.567	0.685	0.566
Lead Inventor Experience	0.103***	0.0330**	0.0304*	-0.008	-0.193**	-0.0374*	0.269**	-0.045
	0.003	0.015	0.076	0.596	0.036	0.083	0.040	0.200
Patent Originality	0.384	0.389	0.417	0.267	-1.486	-0.101	0.299	-0.617
	0.527	0.150	0.215	0.441	0.297	0.830	0.712	0.266
Device	-1.014**	-1.013***	-0.973***	-1.181***	-0.444	-0.210	-39.300	-0.775*
	0.029	0.000	0.000	0.000	0.684	0.454	1.000	0.082
Patent Priority Date in 1981-85	1.438*	1.873***	1.787***	1.644***	1.010	-0.316	-19.160	-0.946
	0.080	0.000	0.000	0.001	0.689	0.542		0.253
Patent Priority Date in 1986-90	1.743***	1.471***	1.436***	1.191**	-0.864	-0.212	-19.67***	-1.096
	0.005	0.000	0.001	0.014	0.692	0.624	0.000	0.108
Patent Priority Date in 1991-95	2.239***	1.518***	1.559***	1.136**	-2.809	-0.831*	-18.09***	-0.898
	0.000	0.000	0.000	0.017	0.196	0.064	0.000	0.183
Patent Priority Date in 1996-00	0.401	1.110***	1.092**	1.371***	2.358	0.077	-20.93***	-0.617
	0.493	0.006	0.014	0.004	0.247	0.862	0.000	0.365
IPC class dummies	yes	yes	yes	yes	yes	yes	yes	yes
Constant	-5.715***	-3.190***	-3.251***	-1.472*	7.626**			
	0.000	0.000	0.000	0.073	0.033			
lnalpha					-0.893***			
					0.002			
Observations	285	285	285	285	285	152	152	152
Groups								62
*** p<0.01, ** p<0.05, * p<0.1								

Table 5a: This table shows logit models with a dependent variable equal to 1 if the confidentiality agreement became a license (i.e. "deal") and 0 if the agreement did not result in a license (i.e. "no deal"). Each observation corresponds to a patent-agreement pair - a patent can have multiple agreements and each agreement can be associated with multiple patents. Patent level measures correspond to the hospital patent which is under the agreement. **Firms for which a cosine measure was not defined because the firm has no patents of its own were assigned a cosine measure of 0 in these models**. See Appendix A for definitions of variables. P-values under coefficients. Robust Standard Errors

	Model 4.1	Model 4.2	Model 4.3	Model 4.4	Model 4.5	Model 4.6	Model 4.7	Model 4.8
Cosine Subclass Level	1.157***	1.285***	1.127***	1.977***	1.361***	1.238***	1.085**	1.276***
	0.006	0.002	0.008	0.001	0.003	0.008	0.022	0.008
Within Section Cosine, Group Level	-0.921**	-1.056**	-0.968**	-2.079***	-1.259***	-1.031**	-0.856*	-1.042**
	0.032	0.014	0.028	0.001	0.010	0.037	0.094	0.041
Patent Scope			0.050	0.051	0.112***	0.105***	0.271***	0.260***
			0.122	0.229	0.005	0.010	0.000	0.000
Number of Patents Cited			0.001	0.009	0.006	0.004	-0.004	-0.006
			0.948	0.375	0.526	0.637	0.763	0.643
Patent Citations Received Per Year			0.373***	0.218**	0.404***	0.378***	0.264**	0.272**
			0.001	0.043	0.000	0.001	0.037	0.033
Patent Originality				-0.195	1.014**	0.853	1.310**	1.291**
				0.717	0.047	0.101	0.028	0.032
Device					-0.332	-0.308	-0.180	-0.110
					0.274	0.321	0.612	0.765
IPCs in Section A Only					1.539***	1.492**	1.746***	1.570**
					0.009	0.012	0.009	0.018
IPCs in Section C Only					2.551***	2.454***	3.264***	3.199***
					0.000	0.000	0.000	0.000
IPCs in Section G Only					0.768	0.828	1.588*	1.519*
					0.344	0.286	0.055	0.076
IPCs in Sections A and C Only					1.573***	1.408**	2.133***	1.923***
					0.008	0.017	0.001	0.004
IPCs in Sections A and G Only					1.948***	1.737**	2.096**	1.889**
					0.005	0.013	0.013	0.020
IPCs in Sections A, C and G Only					0.266	-0.134	-0.552	-0.605
					0.701	0.852	0.510	0.464
Technology Age in Years						0.0760***	-0.005	-0.006
						0.001	0.837	0.812
Patent Priority Date in 1981-85							1.888***	2.038***
							0.004	0.002
Patent Priority Date in 1986-90							2.088***	2.122***
							0.000	0.000
Patent Priority Date in 1991-95							2.762***	2.793***
							0.000	0.000
Patent Priority Date in 1996-00							-0.075	-0.063
							0.884	0.906
Lead Inventor Experience	-0.022	-0.015	-0.019	-0.033	-0.017	-0.019	-0.037	-0.027
	0.204	0.413	0.297	0.132	0.398	0.340	0.120	0.263
Number of Firm Patents		0.000156*						0.000265**
		0.065						0.013
Number of Firm Patents Squared		0.000						-1.78e-08**
		0.557						0.030
Patent Generality				-0.354				
				0.428				
Constant	0.110	-0.131	-0.325	0.347	-2.805***	-2.965***	-4.666***	-4.709***
	0.526	0.484	0.170	0.586	0.000	0.000	0.000	0.000
Observations	588	588	588	440	588	588	588	588

*** p<0.01, ** p<0.05, * p<0.1

Table 5b: This table shows logit models with a dependent variable equal to 1 if the confidentiality agreement became a license (i.e. "deal") and 0 if the agreement did not result in a license (i.e. "no deal"). Each observation corresponds to a patent-agreement pair - a patent can have multiple agreements and each agreement can be associated with multiple patents. Patent level measures correspond to the hospital patent which is under the agreement. **Firms for which a cosine measure was not defined because the firm has no patents are excluded from these models.** See Appendix A for definitions of variables. P-values under coefficients. Robust Standard Errors

	Model 5.1	Model 5.2	Model 5.3	Model 5.4	Model 5.5	Model 5.6	Model 5.7	Model 5.8
Cosine Subclass Level	1.039**	1.430***	0.928**	1.905***	1.434***	1.149**	1.234**	1.683***
	0.019	0.001	0.040	0.003	0.009	0.040	0.034	0.006
Within Section Cosine, Group Level	-0.983**	-1.033**	-1.046**	-2.287***	-1.508***	-1.193**	-1.211**	-1.375**
	0.023	0.017	0.019	0.000	0.006	0.034	0.042	0.022
Patent Scope			0.0694*	0.054	0.108**	0.0879*	0.324***	0.308***
			0.057	0.293	0.019	0.063	0.000	0.000
Number of Patents Cited			-0.003	0.009	0.012	0.008	-0.008	-0.010
			0.778	0.436	0.282	0.486	0.599	0.525
Patent Citations Received Per Year			0.440***	0.387**	0.541***	0.492***	0.356**	0.383**
			0.001	0.016	0.000	0.000	0.025	0.019
Patent Originality				0.729	1.492**	1.199*	1.755**	1.902**
				0.327	0.018	0.064	0.024	0.019
Device					-0.631	-0.605	-0.343	-0.235
					0.100	0.121	0.408	0.583
IPCs in Section A Only					1.773**	1.682**	2.107**	1.953**
					0.037	0.045	0.024	0.042
IPCs in Section C Only					2.723***	2.596***	3.492***	3.495***
					0.003	0.004	0.000	0.001
IPCs in Section G Only					0.364	0.555	1.529	1.358
					0.783	0.645	0.198	0.288
IPCs in Sections A and C Only					1.854**	1.673**	2.591***	2.374**
					0.031	0.048	0.007	0.016
IPCs in Sections A and G Only					2.922***	2.746***	3.320***	3.135***
					0.004	0.006	0.006	0.009
IPCs in Sections A, C and G Only					0.446	0.102	-0.730	-0.746
					0.653	0.922	0.564	0.562
Technology Age in Years						0.104***	-0.010	-0.015
						0.000	0.778	0.687
Patent Priority Date in 1981-85							1.721**	1.847**
							0.018	0.014
Patent Priority Date in 1986-90							1.819***	1.708**
							0.006	0.011
Patent Priority Date in 1991-95							2.170***	2.148***
							0.000	0.000
Patent Priority Date in 1996-00							-0.747	-0.792
							0.190	0.179
Lead Inventor Experience	-0.016	-0.008	-0.006	-0.009	-0.008	-0.005	-0.016	-0.005
	0.432	0.727	0.785	0.754	0.748	0.855	0.607	0.881
Number of Firm Patents		0.000184**						0.000268**
		0.039						0.019
Number of Firm Patents Squared		0.000						-1.70e-08**
		0.418						0.047
Patent Generality				-0.874*				
				0.0913				
Constant	0.168	-0.304	-0.337	-0.233	-3.500***	-3.598***	-5.053***	-5.342***
	0.435	0.226	0.250	0.766	0.001	0.001	0.000	0.000
Observations	424	424	424	304	424	424	424	424

*** p<0.01, ** p<0.05, * p<0.1

Table 6: This table shows logit models with a dependent variable equal to 1 if the FIRST confidentiality agreement became a license (i.e. "deal") and 0 if the agreement did not result in a license (i.e. "no deal"). Each observation corresponds to a patent-agreement pair. **Only the first agreement for each patent was selected.** Patent level measures correspond to the hospital patent which is under the agreement. **Firms for which a cosine measure was not defined because the firm has no patents of its own were assigned a cosine measure of 0 in these models**. See Appendix A for definitions of variables. P-values under coefficients. Robust Standard Errors

	Model 10.1	Model 10.2	Model 10.3	Model 10.4	Model 10.5	Model 10.7	Model 10.7	Model 10.8
Cosine Subclass Level	2.095**	1.465*	1.664*	3.993***	1.847*	1.722	2.725**	2.744**
	0.017	0.090	0.065	0.002	0.090	0.114	0.021	0.021
Within Section Cosine, Group Level	-2.651***	-1.856**	-2.325**	-4.984***	-2.507**	-2.519**	-2.473**	-2.460**
	0.002	0.033	0.011	0.001	0.029	0.022	0.018	0.031
Patent Scope			0.101**	-0.003	0.145*	0.106	0.438**	0.448***
			0.042	0.971	0.051	0.252	0.011	0.009
Number of Patents Cited			-0.001	-0.012	0.011	0.008	0.010	0.012
			0.943	0.608	0.540	0.618	0.731	0.660
Patent Citations Received Per Year			0.299	0.062	0.344	0.350	-0.137	-0.134
			0.186	0.531	0.128	0.113	0.109	0.110
Patent Originality				0.207	0.964	0.935	0.953	1.159
				0.861	0.252	0.287	0.494	0.469
Device					0.097	0.031	0.343	0.380
					0.857	0.957	0.645	0.603
IPCs in Section A Only					2.117**	2.331**	3.986**	3.803**
					0.020	0.031	0.010	0.013
IPCs in Section C Only					3.147***	3.164***	6.434***	6.290***
					0.002	0.006	0.003	0.002
IPCs in Section G Only					1.287	1.287	4.725**	4.085**
					0.258	0.311	0.016	0.022
IPCs in Sections A and C Only					2.074**	2.057*	4.641**	4.441**
					0.022	0.053	0.015	0.018
IPCs in Sections A and G Only					1.924	1.993	3.914	4.154
					0.135	0.154	0.136	0.146
IPCs in Sections A, C and G Only					2.584**	2.772**	4.991**	6.382*
					0.028	0.035	0.023	0.054
Technology Age in Years						0.161**	-0.058	-0.132
						0.021	0.588	0.410
Patent Priority Date in 1981-85							3.338**	3.556**
							0.027	0.039
Patent Priority Date in 1986-90							5.322***	5.178***
							0.000	0.000
Patent Priority Date in 1991-95							4.809***	4.518***
							0.000	0.001
Patent Priority Date in 1996-00							-0.249	-0.300
							0.788	0.760
Lead Inventor Experience			-0.026	-0.0989**	-0.024	-0.024	-0.065	-0.064
			0.436	0.038	0.518	0.528	0.233	0.258
Number of Firm Patents		-0.000661**						-0.001
		0.010						0.306
Number of Firm Patents Squared		6.67e-08***						0.000
		0.004						0.272
Patent Generality				-3.808***				
				0.003				
Constant	0.576***	0.515**	0.267	4.183**	-2.811**	-3.138**	-7.631***	-7.324**
	0.006	0.022	0.480	0.014	0.026	0.035	0.010	0.014
Observations	194	194	194	145	194	194	194	194

*** p<0.01, ** p<0.05, * p<0.1