

Fier, Heide; Pyka, Andreas

Working Paper

Against the one-way-street: Analyzing knowledge transfer from industry to science

ZEW Discussion Papers, No. 10-048

Provided in Cooperation with:

ZEW - Leibniz Centre for European Economic Research

Suggested Citation: Fier, Heide; Pyka, Andreas (2010) : Against the one-way-street: Analyzing knowledge transfer from industry to science, ZEW Discussion Papers, No. 10-048, Zentrum für Europäische Wirtschaftsforschung (ZEW), Mannheim

This Version is available at:

<https://hdl.handle.net/10419/37076>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Discussion Paper No. 10-048

**Against the One-way-street:
Analyzing Knowledge Transfer
from Industry to Science**

Heide Fier and Andreas Pyka

ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 10-048

**Against the One-way-street:
Analyzing Knowledge Transfer
from Industry to Science**

Heide Fier and Andreas Pyka

Download this ZEW Discussion Paper from our ftp server:

<ftp://ftp.zew.de/pub/zew-docs/dp/dp10048.pdf>

Die Discussion Papers dienen einer möglichst schnellen Verbreitung von
neueren Forschungsarbeiten des ZEW. Die Beiträge liegen in alleiniger Verantwortung
der Autoren und stellen nicht notwendigerweise die Meinung des ZEW dar.

Discussion Papers are intended to make results of ZEW research promptly available to other
economists in order to encourage discussion and suggestions for revisions. The authors are solely
responsible for the contents which do not necessarily represent the opinion of the ZEW.

Non technical summary

The expansion of higher education has lead to a continuously increasing number of jobs that deal with scientific problems and methods (Gibbons 1994). In science based industries like the biotechnology industry knowledge has become the most important production input. Since knowledge as a productive factor has different properties than a scarce production factor new economic concepts have been introduced within the last two decades. These new concepts emphasize the interactions between organizations which permanently produce and absorb knowledge so that knowledge flows occur between all actors in all directions. As a consequence knowledge should not only flow between firms or from public research organizations to firms but also from firms to public research organizations. The empirical literature has mostly neglected this topic so far.

This paper analyzes differences in the factors that influence the occurrence of knowledge flows within industry and from industry to science in the biotechnology sector. The knowledge flows are modeled via a backward patent citation analysis on the basis of EPO patent data. We then use an quasi-experimental framework that compares the identified citing and cited patents with a combined sample of control patents. On basis of this combined sample we estimate a weighted bivariate probit model on the citation probability of science and industry. We find considerable differences in the citation probability of science and industry. Cultural closeness has a positive effect on the citation probability from industry to industry while the citation probability of scientific institutions is not affected by cultural distance. Moreover many inventions in the biotechnology sector that are protected by patents obviously seem to be not profitable at a first glance but feature great value for future scientific research because the economic value has only a positive effect on the citation probability of industry. Co-operation between firms and research institutions on a patent application seems to have a signal effect for other research institutions regarding the potential usefulness for own research and thus results in a higher citation rate from science.

Our results suggest that knowledge transfer in the biotechnology industries indeed is not a one-way street between public research organizations and firms but works in both directions. This result qualifies present-day biotechnology industries as science-based industries par excellence as the division of labor in research activities between firms and public research organizations blurs the ancestral boundaries between applied and basic research.

Das Wichtigste in Kürze

In wissenschaftsbasierten Industrien wie der Biotechnologie stellt technologisches Wissen den wichtigsten Produktionsfaktor dar. Technologisches Wissen unterscheidet sich aber in seinen Eigenschaften von knappen Produktionsfaktoren wie Kapital und Arbeit. Aus diesem Grund sind in den letzten beiden Jahrzehnten verschiedene, neuartige ökonomische Konzepte vorgestellt worden. Diese neueren, ökonomischen Konzepte betonen die Interaktion zwischen Organisationen, die auf der einen Seite neues technologisches Wissen produzieren und auf der anderen Seite technologisches Wissen von außerhalb absorbieren. Technologisches Wissen fließt demnach nicht nur zwischen Unternehmen oder von öffentlichen Forschungseinrichtungen zu Unternehmen, sondern sollte auch von Unternehmen zu öffentlichen Forschungseinrichtungen fließen. Die empirische Literatur hat dieses Thema jedoch bislang fast gänzlich ausgeklammert.

Das vorliegende Papier analysiert Unterschiede der Zitationswahrscheinlichkeit zwischen Unternehmen seinerseits und von Unternehmen zu öffentlichen Forschungseinrichtungen in der biotechnologischen Industrie. Die Zitationen von Unternehmen zu Unternehmen und von Unternehmen zu wissenschaftlichen Einrichtungen werden dabei anhand einer Zitationsanalyse auf Basis von EPO-Patentdaten abgebildet. Anhand der dabei identifizierten Patente und dazu gespielter Kontrollpatente schätzen wir ein gewichtetes bivariates Probitmodell, um Unterschiede in den Zitationswahrscheinlichkeiten aufzudecken.

Die empirischen Ergebnisse zeigen, dass Wissenstransfer von Unternehmen zu öffentlichen Forschungseinrichtungen in der biotechnologischen Industrie tatsächlich stattfindet. Dabei gibt es eine Reihe von Faktoren, die Unterschiede in den Zitationsstrukturen von Unternehmen und öffentlichen Forschungseinrichtungen erklären. So haben beispielsweise die kulturelle bzw. räumliche Nähe zwischen zwei Akteuren und der wirtschaftliche Nutzen der patentierten Erfindung einen positiven Effekt auf die Zitationswahrscheinlichkeit von Unternehmen, jedoch keinen Einfluss auf die Zitationswahrscheinlichkeit von staatlichen Forschungsinstituten oder Universitäten. Dahingegen geht von einer gemeinsamen Patentanmeldung von Unternehmen und wissenschaftlichen Einrichtungen ein Signaleffekt für andere wissenschaftliche Einrichtungen aus und erhöht die Wahrscheinlichkeit, dass dieses Patent von wissenschaftlichen Einrichtungen zitiert wird.

Against the one-way-street: Analyzing knowledge transfer from industry to science

Heide Fier^{a1} and Andreas Pyka^b

^a ZEW Centre for European Economic Research, Mannheim (Germany)

^b University of Hohenheim, Hohenheim (Germany)

Juni 2010

Abstract

This study aims at analyzing the differences in the factors that influence the probability of knowledge transfer within industry and from industry to science in the biotechnology sector. In order to model these knowledge flows a citation analysis on the basis of patent data was conducted and a weighted bivariate probit model was estimated on the citation probability of industry and science on the basis of a combined sample of citing and cited patent pairs and an equal number of control patent pairs. The empirical results suggest that there are considerable differences in the citation probability. Cultural closeness for instance has a positive effect on the citation probability from industry to industry while the citation probability of scientific institutions is not affected by cultural distance.

Keywords: Technology transfer, patent citation analysis, biotechnology industry

JEL classification: J61, O33

Acknowledgements:

We thank Thorsten Doherr (ZEW), Patrick Beschorner (ZEW), Christoph Grimpe (ZEW), Georg Licht (ZEW) and the participants at the EMAEE 2009 in Jena for their helpful comments.

¹ Corresponding author, ZEW Centre for European Economic Research, P.O. Box 10 34 43, D-68034 Mannheim (Germany), e-mail: heide.fier@zew.de, phone: +49-228-3728944, fax: +49-621-1235-170

1. Introduction

The expansion of higher education has lead not only to the fact that many people nowadays have acquired substantial knowledge about recent scientific discoveries and research topics but also resulted in a continuously increasing number of jobs that deal with scientific problems and methods (Gibbons 1994). In science based industries like the biotechnology industry science and knowledge has even become the most important production input. However knowledge differs from scarce production factors since it can be “sticky” (von Hippel 1994) which means that knowledge is sometimes so specialized that it can not be easily transferred from one actor to another. In order to capture these preconditions different economic concepts have been introduced within the last two decades which seem to be more suitable for explaining technological change in science based industries compared to neoclassical concepts of scarce recourse allocation. Almost all new economic concepts put knowledge in the middle of their analysis and describe innovative processes as a result of interactions between organizations that permanently produce and absorb knowledge. The concept of innovation systems focuses on the flow of technology between various actors like firms, universities and the government and analyses these technology flows on a regional, national or supranational level (Lundvall 1992, Nelson 1993,). Etzkowitz and Leydesdorff (1997) describe a triple Helix of university-industry-government relations with – in contrast to the national innovation systems approach – a reorganizational component across institutions and national boundaries. One of the most recent concepts from firm theory puts its main focus on the participation in external networks of organization. The idea of “open innovation” was first introduced by Henry Chesbrough (2003) who conducted a number of company based case studies and came to the result that organizations (i.e. firms) have to open themselves up to external networks in order to gain new knowledge. This external knowledge can then be combined with the already existing firm knowledge and capacities for innovative activities can be successfully be build up².

On basis of the introduced economic concepts it becomes visible that knowledge should not only flow from universities or other public research institutions to firms but also vice versa.

² A brief overview of economic concepts that put knowledge in the middle of their perspective can be found in Dogson et al. (2006), pp. 334-335.

However the empirical literature on technology transfer in science based industries has mostly dealt with the question how firms can profit from the research results of scientific institutions or other firms but has disregarded the fact that firms themselves can act as valuable knowledge producers and thus can produce knowledge flows to other firms or scientific institutions.

This paper analyzes differences in the factors that influence the occurrence of knowledge flows within industry and from industry to science in the biotechnology sector. The knowledge flows are thereby modeled via a backward patent citation analysis on the basis of EPO patent data. As a result we are able to identify cited and citing patent pairs. We then use an quasi-experimental design which has been first introduced by Jaffe et al. (1993). This quasi-experimental framework compares the identified citing and cited patents with a matched sample of control patents. On the basis of this combined sample we estimate a weighted bivariate probit model on the citation probability of science and industry.

The structure of the paper is as follows. In the second section we provide a short overview on the characteristics of the biotech industry and the importance of patent protection for this industry. The third section discusses the question whether there is a rationale for technology transfer from industry to science. The fourth section reviews the empirical literature on patent based studies of knowledge flows. The subsequent three sections contain the empirical part of the paper. First the data and methodology are presented (fifth section) and then the variables and descriptive statistics are shown (sixth section). The estimation strategy and the results are presented in the seventh section. Section eight closes with a conclusion

2. Characteristics of the biotech industry and patent protection

As other science based industries the biotechnology industry differs from existing non-science based industries in its pattern and dynamics of technological change. Pavitt (1984) analyses sectoral patterns of technical change by classifying firms according to three dimensions. According to this taxonomy *supplier dominated* firms are characterized by weak R&D and engineering capabilities and their main technology lies in cutting costs through embodied technical change. Thus supplier dominated firms apply rarely for patents. Whereas *production intensive* firms exploit scale economies of production and therefore aim at realizing performance increasing product and process innovations. Product innovations are often protected by patents while process innovations are kept secret. In contrast to the first two

groups *science based* firms depend on the progress of the relevant sciences and their main technology stems from R&D activities of the firms in the sector. Innovations are protected through patents, lead-time advantage and secrecy. However science based industries are not a homogeneous group but include mature industries as well as young industries, and also the R&D intensity varies widely within the science based industries.

Within the science based industries the biotech industry is considered to be a rather young industry and distinguishes itself due to its high R&D intensity (Niosi 2000). The invention of the recombinant DNA technique by Cohen, Boyer and Berg at beginning of the 1970's is often considered to be the starting point for the so called modern biotechnology. Zucker and Darby (1996) were among the first who analyzed the success factors for the formation of the biotechnology industry. In their work they emphasized the role of individual star scientists as a knowledge source for biotech firms. Today the biotechnology industry in developed countries is characterized by a large share of small and medium sized firms which are highly R&D intensive and attract a large amount of money from public subsidiary programs and venture capital agencies (Fuchs 2003). Furthermore biotechnology firms are increasingly producing scientific publications. Gittelman and Kogut (2003) have analyzed a sample of 116 US biotech firms in the time period between 1988 and 1995. They show that the total publication rate of the firms almost doubled in that time span.

With the rise of the modern biotechnological industry and the growing awareness of the economic and sociological potential of this industry a major problem occurred in how the intellectual property of biotechnological inventions could be protected. The existing patent protection laws in the US and other countries at the beginning of the 1980s were not designed for the protection of biotechnological inventions. With a broadened definition of patentable subject matters due to a change of the patent protection law in the US in 1992 and subsequent changes of patent protection laws in other countries it became possible to protect biological active substances including single molecules and proteins (Ko 1992). Therefore patents create a basis for trading inventions. As a consequence patents have great importance in biotechnology not only in the protection of marketable inventions and thus as a positive signal for venture capital firms but also for discoveries that are not marketable at the first glance but feature great value for further research (Mazzoleni and Nelson 1998).

3. Knowledge interactions in the biotechnology industry – is there a rationale for knowledge transfer from industry to science?

Knowledge is nowadays considered to be an indispensable factor for economic growth. Arrow was the first who stressed the importance of knowledge for economic growth. In his model, Arrow assumes that new knowledge is created depending of the level of new investments and in turn the technologies accessible for firms depend on the economy wide knowledge stock. This Arrowian view suggests that technological knowledge has the non-excludable and non-rival character of a public good and can be transferred and appropriated with rather low effort and costs (Arrow 1962, Arrow 1969).

This rather traditional approach to the nature of technological knowledge has been challenged by the Neo-Schumpeterian approach in recent years. In the view of Neo-Schumpeterians technological knowledge is considered to be a quasi-public good, which means that the character of technological knowledge bears higher levels of appropriability and excludability compared to the Arrowian view (Rosenberg 1994, Antonelli 1999). Moreover the production of technological knowledge is considered to be path-dependent and cumulative and can have a local character. This Neo-Schumpeterian view of technological knowledge implies that “knowledge is the result of complex processes of creation of new information building upon the mix of competences acquired by means of learning processes, the socialisation of experience, the recombination of available information and formal R&D activities” (Antonelli 1999, p. 245). The innovation system approach confirms this view and emphasizes the importance of interactions between industry and science for a successful innovation process due to its increasing complexity (Lundvall 1992, Nelson 1993, Nelson and Rosenberg 1993). Moreover, a number of studies have examined the relationship between the technological complexity (measured by the R&D intensity) of industries and the number of R&D alliances and they have found a positive correlation between these two factors (e.g. Freeman 1991, Hagedoorn 1995).

Owing to their science based nature, problems of appropriability and excludability of technological knowledge are even more severe in the modern biotechnological industry (Arora and Gambardella 1990). In order to succeed in the biotechnology industry firms must permanently keep close contact to the moving technological frontier and must create valuable technological knowledge on their own (Gambardella 1995, Niosi 2003). Thus, the ability of firms to draw knowledge from scientific institutions or other firms is regarded to be an

important factor for their success (Kenney 1986, Prevezer 2001, Niosi 2003, Powell et al. 1996). Several studies have shown that geographical closeness between firms and research institutions in biotechnology can facilitate this knowledge transfer from science to industry (Zucker et al. 1994, Audretsch and Stephan 1996, Zucker and Darby 1998, Powell et al. 1999). Also the role of individual scientists for the prosperity of firms in biotechnology has been highlighted (Zucker and Darby 1996).

Besides the critical role of knowledge flows from science to industry also knowledge interactions between firms in biotechnology have been recognized to be crucial for the industrial development. Pyka and Saviotti (2005) analyze research networks in the biotechnology industry and conclude that a coexistence of large diversified firms and small dedicated biotech firms is crucial for industrial development. For small firms a co-operation with large pharmaceutical or chemical firms can result in the gain of more market relevant knowledge in the form of the use of advanced production capabilities, better market access due to a better distribution infrastructure and experience in conducting clinical trials (Pisano 1990, Baum et al. 2000). In turn large firms in the biotech sector seek to co-operate with small/medium sized research intensive firms in order to acquire marketable knowledge and to spread risks (Arora and Gambardella 1990).

Whilst these two directions of knowledge flows namely from science to industry and within industry have been fairly well analyzed there is a lack of evidence regarding knowledge flows from industry to science. The main reason for the negligence of research on knowledge transfer from industry to science is the threat of a negative influence of technology transfer upon the norms of open science (Merton 1973). In traditional sectors like manufacturing, universities and public research institutes are still considered to be the most important producers of valuable scientific research (Gibbons et al. 1994).

In the biotechnology industry however, things look different. Due to the mentioned science base of the industry, firms themselves next to public research organizations have accumulated a large stock of technological knowledge. This creation of technological knowledge within firms has been accelerated by venture capital firms with the aim of realizing returns due to groundbreaking inventions as well as public subsidy programs with the objective of not falling behind the industrial development compared to other countries. As a result, there is a considerable amount of valuable technological knowledge in the biotechnology industry that has not been transferred by research institutions in the first place but has been created within

the firms. Pisano (1990) conducted an empirical analysis among US firms and found that firms in biotechnology rely more often only on technological knowledge which has been created in-house especially in those areas where the firms have accumulated in-house R&D.

Thus the question arises whether public research organizations in the field of biotechnology are willing and able to participate in the knowledge that has been produced by firms. The existing literature on this topic is rare however and there are no specific studies for the biotechnology industry. Meyer-Krahmer and Schmoch (1998) have conducted a survey among professors from universities or public research institutions in science based fields and asked them to rate the importance of different interaction types with industry. As a result the interviewed professors rated those interaction types with industry higher where a bidirectional exchange of knowledge with industry occurs. Link et al. (2007) have examined knowledge transfer between industry and science on the basis of a survey among individual scientists. Their results suggest that university researchers rank collaboration with industry as very important and state that they benefit from the transferred knowledge and the use of enhanced equipment. Kaufman and Tödtling (2001) emphasize the importance of a bidirectional knowledge exchange between industry and science in innovation co-operations. It becomes obvious that knowledge transfer from industry to science has not been completely neglected in previous studies but it is mostly mentioned as a by-product from science to industry knowledge flows. This study aims at contributing more empirical evidence to the topic of industry to science knowledge flows in the biotechnology industry.

4. Review on patent based studies of knowledge flows

Patent data have been extensively used to shed light on the innovation process. Patent documents provide information about the technology of an invention as well as detailed information about the inventor and assignee of the invention. For example patent counts have been frequently used as an indicator of innovation activity. However, patent data should be handled with some caution. Griliches has surveyed in his seminal work the pitfalls that may arise when using patent statistics as innovation indicators but concludes that “Nothing else even comes close in the quantity of available data, accessibility, and the potential industrial, organizational, and technological detail” (Griliches 1990, p. 1702).

The idea to use patent data as an indicator for knowledge flows can be traced back to Schmookler (1966) and Scherer (1982). Schmookler among others brought up the discussion,

that the economic benefits of firms due to R&D could not be solely reduced to their own R&D activities, but also to the embodiment of technological knowledge through intermediate products produced by other sectors. Scherer (1982) took up Schmooklers idea and developed a complex “interindustry technology flows” matrix which traces back the knowledge of R&D performing industries to industries that purchased the products of the R&D performing industries. In a following work Scherer relied on a data set that contains over 15.000 US patents that were individually examined to determine the original industry of the patent and the industries for which the use of the patent was anticipated and linked them to the R&D outlay of corporations. The linked R&D outlays were then distributed through a “technology flows” matrix. The estimation results indeed revealed the critical role of embodied technological knowledge for firms’ productivity growth (Scherer 1982, Griliches and Lichtenberg 1984).

More recent work that use patent citations to trace knowledge flows mostly deal with the question whether knowledge flows are technologically bounded, geographically concentrated and what industry specific differences exist (e.g. Jaffe et al. 1993, Jaffe and Trajtenberg 1996, Porter 2000, Maurseth and Verspagen 2002).

Stolpe (2002) modeled the citation probability among patents in the liquid crystal display technology and revealed that technological closeness has a significantly positive influence on the citation probability. However Stolpe (2002) did not make a distinction between the institutional types of the assignees of the citing patents. Hu and Jaffe (2003) have worked out the positive effect of technological closeness for the citation probability in a cross country comparison. Besides the technological closeness also the technological generality of the cited patent may have an influence on the citation probability. Trajtenberg et al. (1992) have shown that university research outcomes are more basic and harder to appropriate than research outcomes of industry.

The hypothesis that knowledge flows might be geographically bounded has been heavily analyzed and discussed within the last years. Firms that have the same cultural background are more likely to exchange knowledge than firms with different cultural backgrounds. Mowery et al. (1996) have shown that more knowledge exchange takes place in alliances with partners who have the same nationality. Empirical evidence is less clear regarding geographical closeness. Jaffe et al. (1993) were the first who found direct evidence that knowledge spillovers as measured by patent citations are indeed locally concentrated.

Although the quasi-experimental design that was used by Jaffe et al. (1993) was challenged afterwards (Thompson and Fox-Kean 2005, Thompson 2006), the empirical evidence could not be disproved. Although doubts remain from the theoretical perspective (Breschi and Lissoni 2001) it is supposed that geographical closeness has a positive impact on the citation probability.

A few recent studies have analyzed knowledge flows in the biotechnological sector on the basis of patent data. McMillan et al. (2000) have worked out the importance of public science for the development of the US biotechnology industry on the basis of patent data. The authors conclude that especially small biotech firms depend on the basic knowledge that is created by public research organizations. Gittelman (2006) has examined the differences in the public-private knowledge flows between the US and France on the basis of patent citations. In line with the work of Zucker and Darby (1998) emphasize the importance of individual scientific careers for interactions between firms and public research organizations. Moreover they point out that technological performance, as measured by the number of granted patents, depends on a heterogeneous setting of organizations and interactions.

5. Data and methodology

Patent citation analysis and data

The aim of the study is to analyze the differences in the factors that influence the probability of knowledge transfer within industry and from industry to science in the biotechnology sector. In order to model these knowledge flows we conduct a backward patent citation analysis: for each patent in the sample all citations which have been made by timely subsequent patents in the sample are identified.

The study is based on patent application data from the European Patent office (EPO)³ which cover the years between 1978 and 2003. The patent data include information about the name(s) and country(ies) of origin of the inventor(s) as well as the assignee(s), the declared IPC classes as well as application and grant dates. Moreover a patent document contains references to other patents, so called citations. In EPO patent data, these citations have mainly

³ For a comprehensive overview on the application and examination process at the EPO see Michel and Bettels (2001). For differences in the examination process at the EPO and other patent offices see

the legal function to specify the knowledge that justifies a claim for novelty and are mostly added by the patent examiners instead of the inventors. Alcácar and Gittelman (2006) find that examiners played a significant role in identifying prior art, adding 63% of citations on the average patent, and all citations on 40% of patents granted. This might be due to two reasons. Either the inventors are not aware of the patents that have been added by the examiners (Criscuolo and Verspagen 2008) or the inventors have strategically omitted citations (Alcácar and Gittelman 2006). Nevertheless, since we focus on a rather small technological field the actors should not have problems identifying prior art (Maurseth and Verspagen 2002). Regarding strategic omission of prior art the patent examiners and the application process of the EPO plays an important role. In the patent application process, the applicant receives a detailed search report, conducted by the patent examiners, which discloses essential prior art on which the examiner would mainly base his grant decision. After obtaining the search report, the applicant must decide whether he wants to pursue the application process or not. Thus the risk that prior art remains undetected is minimized by the work of the patent examiner.

In a first step we identify on the basis of the OECD compendium of patent statistics (OECD 2008) all relevant international patent classification (IPC) classes concerning biotechnology. Following this classification scheme all records where at least one of the relevant IPC classes was listed in the application are kept for further analysis. Subsequently all applicants in the data files are assigned by hand to the following categories: firms, universities, public research institutions, individuals, others. Our data cover 72427 patents that have been applied for in the mentioned time period. We use the application date of the patent application as the relevant time point for our analysis as it is common in most of scientific works that deal with patent analysis.

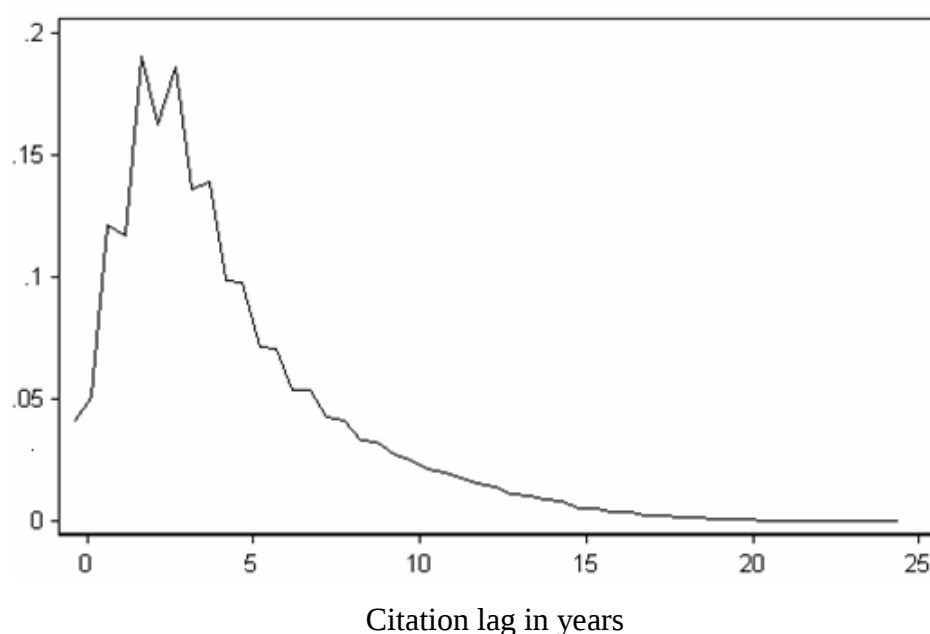
Truncation and restriction of the sample

Since the analysis concentrates on a comparison of the knowledge flows from industry to science and within industry only those patent pairs were kept, where at least one firm was among the applicants of the cited patent and at least one firm and/or one research institution was among the applicants of the citing patent. However due to the fact that in many countries scientists had or still have the privilege to assign patents under their own name, the share of scientific applicants is likely to be underrepresented.

Moreover all patent pairs where the cited patent and citing patent showed the same application name, so called self citations were excluded from the sample since they solely reflect in-house knowledge flows.

The application of a patent citation approach necessitates considering one difficulty, because the patent that has been filed first in the sample has a much larger time frame to be cited compared to the patent that has been filed more recently in the sample. This problem of truncation has been heavily analyzed in empirical studies. Caballero and Jaffe (1993) and Jaffe and Trajtenberg (1999) estimated the shape of the citation lag distribution via a parametric function and Hall et al. (2001) used non-linear functions to approximate the shape via estimation. Stolpe (2002) states in his work that in the ideal case citation studies should be based on patents that have been filed at exactly the same point in time so that the problem of temporal influences on the citation frequency can be neglected. However in the same breath he accounts for the fact that patent data are flow data and that they are thus only measurable over time. In his study he sets a time limit of three years for the selection of the patents that are later referred to via citation analysis. Almeida (1996) deals with the problem of truncation by including the citation lag in his latter estimation. Gittelman (2006) includes not only the citation lag but also the square of the citation lag in her regression and moreover limits the time span where the cited patents are identified.

Figure 1: Citation lag – Kernel density estimation



Within our sample, the mean citation lag is 4.2 years. Figure 1 shows the distribution of the citation lag by means of a Kernel density estimation. The cited patents receive most of their citations in the second and third year after their application date. Only 10 percent of the citations have been made more than 9 years after the EPO application date of the cited patent. Thus by following the approach of Gittelman (2006) and including not only the citation lag but also the squared citation lag in the following estimation it is accounted for the fact that there are rather few patents with a very short or very long citation lag. Since we have patent data available from 1978-2003 and the mean citation lag in our sample is 4.2 years we limit the time span in which we select the cited patents to 1978-1998, so that patents which are issued in 1998 have a rather equal chance of being cited.

Construction of a control sample

Because we aim at analyzing differences in the factors that influence the citation probability we need to include reference values to the sample of identified cited and citing patent pairs in order to maintain interpretable results. For this purpose we follow an experimental design which was first introduced by Jaffe et al. (1993) and later used by several other studies (e.g. Almeida and Kogut 1999, Stolpe 2002). Within this experimental framework, a non-cited patent that shows the same first three digit international patent classification (IPC) class and the same EPO application date as the cited patent is randomly searched for each citing patent within the original sample. However it is important to note that the fact that a patent is chosen to be a control patent for a specific citing patent does not mean that it can not have received citations in an earlier or later point of time.

Due to the construction of the control sample we are able to model an unconditional probability for the factors influencing the citation probability. The conditional probability for the influencing factors is given when an actual citation has occurred. Thus the hypothesis that can be tested is whether a statistically significant difference between the conditional and unconditional probabilities exists when examining the citation probability⁴.

⁴ The two probabilities are related. Bayes rule states that $P(\text{Citation} | \text{Influencing factor})/P(\text{Citation})=P(\text{Influencing factor}/\text{Citation})/P(\text{Influencing factor})$

6. Variables and descriptive statistics

The dependent variables in our estimation *INDIND* and *INDSCI* are binary variables indicating whether a patent that has been applied for by industry has received a citation by a patent that was applied for by either industry (*INDIND*) and/or scientific institution(s) (*INDSCI*).

Building upon the previous discussion a set of independent variables was included in the estimation that is likely to have an influence on the citation probability.

First a variable to proxy the technological closeness of the patent pair was included. *TECHCL* is a dummy variable indicating whether the two patents in a patent pair show the same 6-digit IPC class. Since we look at industry research outcomes as possible appropriable targets we assume that their technological character might be less basic. Nevertheless things might turn out to be different for two reasons. First we look at a science intensive industry where a large part of industrial actors are involved in basic research and second we only consider knowledge flows from industry to industry and industry to science. However, we assume that a high technological generality⁵ implies a more basic technological character of the invention of the cited patent and therefore expect that it is positively related with the citation probability from industry to science and on the other side we assume that a more specific technological character of a technological invention is positively related with the citation probability within industry. While previous works that measured the technological specialization of patents on the basis of IPC-classes often used the Herfindahl-index, van Zeebroeck et al. (2006) have compared different technological concentration measures on the basis of EPO patents and come to the result, that the Gini-Coefficient⁶ in line with the C20-measure are the most reliable measures for technological concentration. Moreover they recommend at least a 4-digit aggregation level of the IPC-classes used. Consequently, this study relies on the Gini-Coefficient for the identified biotechnology related IPC classes of the citing patents aggregated to the 6-digit level as a measure for technological specialization and we specify

⁵ Generality is also referred to as basicness. See i.e. Stolpe (2002).

⁶ The Gini Coefficient is a statistical measure for relative Concentration. It relies on the concept of the Lorenz Curve. The Gini Coefficient takes on values between zero and one, whereas the value one corresponds to perfect inequality.

generality = 1 - Gini coefficient

as a proxy for the technological generality of the cited invention. In those cases where the Gini-Coefficient is calculated on the basis of only one IPC-class the measure for generality is replaced with zero. However, since the Gini-Coefficient reduces complex data to one parameter, there is the danger that valuable information from the used data is neglected. In this case, the Gini-Coefficient does not account for the number of different IPC classes of the citing patents, although this is obviously valuable information when approximating the technological generality of an invention. In order to account for this shortcoming we include an interaction term between the measure of technological generality and the total number of different IPC-classes of the citing patents (*INTGINI*) in the regression instead of the plain measure of technological generality.

The variable *CULCL* indicates the cultural closeness of the patent pairs. *CULCL* is a dummy variable and measures whether the two patents in the patent pairs have the same assignee countries.

Besides that we include a dummy variable reflecting whether the cited or cited control patent has been assigned by both, industry and science (*COMMON_CITED*). It is important to note that the variable *COMMON_CITED* is a rough indicator for joint research, since firms and research institutions can of course conduct joint research without being jointly listed as assignees in a particular patent application. However, a joint assignment of the common research might signal that the protected invention has a major value for both the scientific and the industrial progress.

Moreover we include variables that reflect the overall patenting activity in the biotechnology field of the assignee(s) of the cited or cited control patent (*NOPATS_CITED*) and the patenting activity of the applicant(s) of the citing patent (*NOPATS_CITING*). *NOPATS_CITED* and *NOPATS_CITING* are continuous variables and contain the cumulated number of patents that the assigning institution(s) have applied for up to the EPO date of the considered patent in the particular patent pair. It is expected that a high patenting activity of the assignee(s) of the citing patent (*NOPATS_CITING*) is positively related to the citation probability, especially regarding scientific institutions as assignees. This assumption is owed to the work of Owen-Smith and Powell (2001) who revealed that scientific institutions which patent have a higher propensity to engage in technology transfer. More precisely they

analyzed the propensity of scientific research institutions transferring knowledge to other firms or research institutions. However we assume that scientific institutions that patent might also show a higher probability to draw knowledge from patents that have been applied for by firms since the scientific institutions might be better informed about the patented inventions of firms due to a review of existing inventions during the application process.

We also include a proxy for the economic value (*ECVALUE*) of the cited respectively cited control patent. *ECVALUE* contains the whole number of subsequent citations that a cited patent has received on the basis of our original sample. Harhoff et al. (1999) obtained value estimates of inventions filed in patents due to a survey of the patent owners. They found a significant positive relationship between the private value estimate of the invention of the filed patent and the number of subsequent citations of this specific patent. Hall et al. (2001) have confirmed this positive relationship. In their work they compare different measures that are likely to influence the market value of firms and conclude that a citation weighted patent stock is more highly correlated with the market value than the plain patent stock. Since we expect firms to be profit oriented we expect that they transfer knowledge from the economically most valuable inventions.

Moreover we also control for the country of residence of the assignees at the time point of their patent application. Since we have a large number of applicant countries in our sample we decide to only include dummies for the three countries that account for most of the patent applications in our sample. As a consequence the included country dummies have to be interpreted in relation to all other countries that are not captured via the country dummies. For example the variable *US_CITED* contains the information whether at least one of the assignees of the cited or cited control patent was located in the United States during the patent application process and *US_CITING* contains the same information for the assignee of the citing patent. Analogous dummy variables were created for Japan (*JP_CITED*, *JP_CITING*) and Germany (*GER_CITED*, *GER_CITING*)⁷.

As already discussed we include two variables to control for the citation lag. *YEAR_DIFF* and *YEAR_DIFFSQ* are continuous variables which reflect the time lag between the cited or cited

⁷ See also the OECD Biotechnology statistics (OECD 2006) for a more general overview on the patenting activities of different countries at the EPO.

control patent and citing patents and the controls, measured by years. We further include dummies for the application year (*YEAR1-YEAR23*) of the cited and cited control patent to control for intertemporal differences in the patenting activity.

Descriptive statistics

Table 1 shows the descriptive statistics of the combined sample. Due to the construction of the sample the control patent pairs account for exactly half of the data. It can be seen that knowledge flows from industry to science are rare but indeed happen. About 10% of the patents that were filed by firms and that received citations received them by public scientific institutions (*INDSCI*).

Table 1: Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
<i>INDSCI</i>	30210	0.053	0.225	0	1
<i>INDIND</i>	30210	0.454	0.498	0	1
<i>TECHCL</i>	30210	0.431	0.495	0	1
<i>CULCL</i>	30210	0.292	0.455	0	1
<i>INTGINI</i>	30210	3.702	2.661	0	16
<i>COMMON_CITED</i>	30210	0.017	0.130	0	1
<i>ECVALUE</i>	30210	9.610	11.125	1	112
<i>NOPATS_CITED</i>	30210	118.119	169.439	1	1083
<i>NOPATS_CITING</i>	30210	62.331	95.131	2	644
<i>YEAR_DIFF</i>	30210	4.277	3.486	0	24
<i>DE_CITED</i>	30210	0.104	0.306	0	1
<i>US_CITED</i>	30210	0.428	0.495	0	1
<i>JP_CITED</i>	30210	0.224	0.417	0	1
<i>DE_CITING</i>	30210	0.125	0.330	0	1
<i>US_CITING</i>	30210	0.398	0.489	0	1
<i>JP_CITING</i>	30210	0.218	0.413	0	1

With respect to the technological closeness (*TECHCL*), we find that about 40 % of the examined patent pairs show the same 6 digit IPC class. Technological closeness can thus be observed more often than cultural closeness (*CULCL*). Only about 30% of the patent pairs show the same assignee country. Due to the fact that the indicator for technological generality of the cited or cited control invention (*INTGINI*) is an interaction term, the interpretation of the descriptive statistics is rather vague. Yet the high standard deviation indicates that the distribution of *INTGINI* is rather unequal.

Only a small number of patents in our sample have been jointly applied for by science and industry. The variable *COMMON_CITED* indicates that not more than 2 % of the inventions

in our sample have assignees from both industry and science. The actual number of joint patent applications between industry and science in the relevant time span is assumingly higher since in many countries scientists had and still have the privilege to freely realize the economic benefits of their inventions⁸. So it is important to bear in mind that *COMMON_CITED* can only be interpreted as a rough indicator for joint research between science and industry. The variable that reflects the economic value of the patented invention shows that on average a cited or cited control patent receives citations from almost 10 other subsequent patents. Regarding the overall patenting activity in the biotechnology field of the assignee(s) of the patent pairs we find that the assignees of the cited patents (*NOPATS_CITED*) have applied for almost twice as many patents as the assignees of the citing patents (*NOPATS_CITING*). Moreover the descriptive statistics show that most of the patents in our sample have been assigned by firms or research institutions from the United States (*US_CITED*, *US_CITING*).

7. Estimation strategy & Results

The focus of this paper is to investigate differences in the citation probability from industry to industry and from industry to science. Thus our two dependent variables in the estimation indicate whether a patent that has been assigned to industry has either received a citation by a scientific institution (*INDSCI*) or by a firm (*INDIND*).

In order to get a first hint on differences between the citation probability of industry and science, we conducted t-tests. The results can be found in the annex table A1. However the t-tests just indicate whether there is a significant difference in the mean values of the variables but can not provide information about the size of these effects. Therefore a discrete probability model is applied.

Because we have two dependent variables and an invention can receive patent citations from both, industry and science a bivariate probit model is estimated. The bivariate probit model estimates the two citation decisions simultaneously and allows the error terms to be correlated.

⁸ I.e. in Germany this privilege was not changed until 2002.

Because we restrict our sample to cited and citing patent pairs and their controls we apply sample weights to the regression to avoid bias from a probability based sample. The sample weights show the probability that a patent pair was chosen from the sample. Thus for patent pairs where the cited patent shows a more recent application date, the probability for a citing patent to be chosen from all possible subsequent patents is higher compared to cited patents with an earlier application date. Additionally, the probability that a cited and citing patent pair was chosen from the sample is lower than the probability that a control patent pair was chosen from the sample. In the weighted bivariate regression the sample weights are included as inverts such that patent pairs with a lower probability to be chosen are weighted higher for the estimation in relation to those patent pairs with a higher sample inclusion probability.

Table 2 shows the estimation results of the weighted bivariate probit model. Tables reporting the marginal effects and the correlation among the variables of the estimated bivariate probit model can be found in the annex table A2 and A3.

Technological closeness (*TECHCL*) of the two patents in a patent pair has a significant positive effect on the probability to be cited from both industry and science, thus the findings of Stolpe (2002) and Hu and Jaffe (2003) are confirmed.

Whereas cultural closeness (*CULCL*) has a positive significant effect on the citation probability of industry, it does not matter for the citation probability of scientific institutions. A possible explanation is that the knowledge flow is highly related to the persons involved in the research process such that spillovers among firms are facilitated from cultural proximity (i.e. Porter 2000a, Mowery et al. 1996). In contrast, researchers from scientific institutions are forced to conduct a thorough search for prior art and related works when writing for academic publications. Therefore they are less likely to be affected by cultural distance.

The interaction term (*INTGINI*) that reflects the technological generality of the cited or cited control invention suggests that firms and scientific institutions are more likely to cite industrial inventions with a less broad technological character. While we expected to find a negative relationship between an increasing technological generality of the possible cited invention and the citation probability of industry we also find that this relationship holds for the citation probability of scientific institutions. Thus this finding might confirm the work of Link et al. (2007) who have examined knowledge transfer between industry and science on

the basis of a survey among individual scientists and come to the result that the scientists can benefit from the applied knowledge of firms and from the use of their enhanced equipment.

The indicator for joint research *COMMON_CITED* shows a highly significant positive effect on the citation probability from industry to science. As pointed out previously the indicator for common research is rather blurry since we expect that more firms have conducted common research with science on the patented invention with the difference that these scientific institutions were not listed as applicants in the patents applications. Still a joint patent application between science and industry obviously signals the scientific relevance of the invention to other research institutions and thus increases the probability of a scientific citation.

Table 2: Results of the weighted bivariate probit model

Variable	Industry to science (INDSCI)		Industry to industry (INDIND)	
	Coef.	Std. Err.	Coef.	Std. Err.
<i>TECHCL</i>	0.120	0.039 ***	0.684	0.023 ***
<i>CULCL</i>	0.054	0.044	0.375	0.027 **
<i>INTGINI</i>	-0.030	0.015 **	-0.017	0.009 ***
<i>COMMON_CITED</i>	0.627	0.090 ***	-0.047	0.067 ***
<i>ECVALUE</i>	0.008	0.007	-0.008	0.004 ***
<i>ECVALUESQ</i>	-0.0001	0.000	0.0001	0.000 **
<i>NOPATS_CITED</i>	0.000	0.000	-0.0001	0.000 ***
<i>NOPATS_CITING</i>	-0.006	0.000 ***	0.001	0.000 **
<i>YEAR_DIFF</i>	-0.008	0.013	0.019	0.008
<i>YEAR_DIFFSQ</i>	0.001	0.001	-0.001	0.001 **
<i>DE_CITED</i>	-0.203	0.068 ***	-0.081	0.041
<i>US_CITED</i>	-0.018	0.052	-0.228	0.030 **
<i>JP_CITED</i>	-0.119	0.049 **	-0.135	0.030 ***
<i>DE_CITING</i>	0.172	0.072 **	0.022	0.039 ***
<i>US_CITING</i>	0.127	0.043 ***	-0.147	0.028
<i>JP_CITING</i>	-0.327	0.060 ***	0.074	0.033 ***
<i>CONS</i>	-1.079	0.104 ***	-0.503	0.068 ***
<i>ATRHO</i>	-0.481	0.027 ***		
<i>RHO</i>	-0.447	0.022		
Wald test of rho=0: $\chi^2(1) = 309.639$ Prob > $\chi^2 = 0.0000$				
30210 observations				

Note: ***, **, * indicate a significance level of 1%, 5%, 10%, Year dummies are included.

The economic value (*ECVALUE*) of the patented invention expressed by the total number of received subsequent citations has no influence for the citation probability of scientific institutions whereas for firms a highly significant u-shaped relationship between the economic value of the cited patent and the citation probability can be found. This u-shaped relationship indicates that contrary to our expectation firms do not only draw knowledge from in an economic sense most valuable inventions but are equally interested in inventions which are

characterized by a comparable low economic value. A possible explanation for the missing significant relationship between the economic value of an invention and the citation probability of science could be the mentioned fact that many inventions that are protected by patents are not marketable at a first glance and thus do not bear a high economic value but are characterized by a considerable value for further scientific research (Mazzoleni and Nelson 1998).

The patenting experience of the applicant firm of the cited patent (*NOPATS_CITED*) has only a positive effect on the citation probability of research institutions. Accordingly, research institutions are more likely to cite patents that have been applied for by firms who are producing a high knowledge output. In contrast to these findings an increase in the accumulated number of patents of the citing (*NOPATS_CITING*) research institutions bears a significantly negative probability for the research institution to cite industry patents. This finding is opposed to our assumption that research institutions that have a high number of accumulated patents might show a higher probability to transfer knowledge from industry patents. Obviously research institutions in the biotechnology sector are screening the knowledge that has been created by firms on a regular basis, especially when they are not frequently patenting. On the contrary assignee firms of the citing patents that show a high patenting activity are more likely to cite patents from other firms. Thus the mentioned necessity for biotechnology firms to acquire external knowledge from other firms to keep up with the technological frontier even when they are actively involved in own research is confirmed by this result.

The included country dummies show opposed effects on the citation probability of industry and science. However the following results have to be interpreted with caution due to differing privileges in the economic usage of inventions of scientists among the countries. Our results show that Germany and Japan as assignee countries of the potentially cited inventions lowers the citation probability of science and the US and Japan as assignee countries of citable inventions have a negative effect on the citation probability of industry compared to the other countries. In contrast to this the citation probability of science increases if the (potentially) citing patents show Germany and the US as assignee countries and the citation probability decreases if the assignee country of the potentially citing patents are assigned by a Japanese scientific institution. In turn citing patents with Germany or Japan as assignee countries have a positive effect on the citation probability of industry. Thus the results suggest that German and US research institutions compared to Japanese and other research institutions

are more actively involved in screening and transferring knowledge that has been produced by industry. As pointed out before and confirmed by these results cultural proximity obviously plays no role for the knowledge transfer from industry to science.

8. Conclusion

This paper aims at investigating differences in the citation probability from industry to industry and from industry to science. We estimated a weighted bivariate probit model on the citation probability of industry and science on the basis of a combined sample of citing and cited patent pairs and an equal number of control patent pairs.

The empirical results suggest that there are considerable differences in the citation probability. Cultural closeness has a positive effect on the citation probability from industry to industry while the citation probability of scientific institutions is not affected by cultural distance. The economic value has only a positive effect on the citation probability of industry but again has no effect on the citation probability of science. However many inventions in the biotechnology sector that are protected by patents obviously seem to be not profitable at a first glance but feature great value for future scientific research. Co-operation between firms and research institutions on a patent application seems to have a signal effect for other research institutions regarding the potential usefulness for own research and thus results in a higher citation rate from science.

Our results suggest that knowledge transfer in the biotechnology industries indeed is not a one-way street between universities and other public research institutions and firms but works in both directions. This result qualifies present-day biotechnology industries as science-based industries par excellence as the division of labor in research activities between firms and public research organizations blurs the ancestral boundaries between applied and basic research.

Literature

Alcázar, J. and M. Gittelman (2006) Patent Citations as a Measure of Knowledge Flows: The Influence of Examiner Citations, *The Review of Economics and Statistics* 88(4), 774-779.

Almeida, P. (1996), Knowledge sourcing by foreign multinationals: Patent citation analysis in the U.S. semiconductor industry, *Strategic Management Journal* 17, 155-165.

Almeida, P. and B. Kogut (1999), Localization of Knowledge and the Mobility of Engineers in Regional Networks, *Management Science* 45(7), 905-917.

Antonelli, C. (1999), The Evolution of the Industrial Organisation of the Production of Knowledge, *Cambridge Journal of Economics* 23(2), 243-60.

Arora, A. and A Gambardella (1990), Internal Knowledge and External Linkages: Theoretical Issues and an Application to Biotechnology, *Journal of Industrial Economics* 38(4), 361-379.

Arrow, K.J. (1962), The Economic Consequences of Learning by Doing, *Review of Economic Studies*, XXIX (3), 155-173.

Arrow, K.J. (1969), Classificatory Notes on the Production and Transmission of Technological Knowledge, *American Economic Review* 59(2), 29-35.

Audretsch, D.B. and P.E. Stephan (1996), Company-Scientist Locational Links: The Case of Biotechnology," *American Economic Review* 86(3), 641-652.

Baum, J.A.C, Calabrese, T. and B.S. Silverman (2000), Don't go it alone: alliance network composition and startups' performance in Canadian biotechnology, *Strategic Management Journal* 21(3), 267-294.

Breschi, S. and F Lissoni (2001), Knowledge Spillovers and Local Innovation Systems: A Critical Survey, *Industrial and Corporate Change* 10(4), 975-1005.

Caballero, R.J. and A.B. Jaffe (1993) How High are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth, *NBER Working Papers* 4370, National Bureau of Economic Research.

Chesbrough, H.W. (2003), *Open innovation: The new imperative of creating and profiting from innovation*, Harvard Business School Press, Boston.

Criscuolo, P. and B. Verspagen (2008), Does it matter where patent citations come from? Inventor vs. examiner citations in European patents, *Research Policy* 37(10), 1892-1908.

Dogson, M., Gann, D. and A. Salter (2006), The role of technology in the shift towards open innovation: The case of Procter & Gamble, *R&D Management* 36(3), 333-346.

Etzkowitz, H. and L. Leydesdorff, Eds. (1997), *Universities in the Global Knowledge Economy: A triple helix of university-industry-government relations*, Cassell, London.

Freeman, C. (1991), Networks of innovators: A synthesis of research issues, *Research Policy* 20(5), 499-514.

Fuchs, G. ed. (2003), *Biotechnology in comparative perspective*, Routledge, London.

Gambardella, A. (1995), *Science and Innovation: The US Pharmaceutical Industry during the 1980s*. Cambridge Univ. Press, Cambridge.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P. and M. Trow (1994), *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*, Sage, London.

Gittelman M, Kogut B. (2003), Does Good Science Lead to Valuable Knowledge? Biotechnology Firms and the Evolutionary Logic of Citation Patterns. *Management Science* 49(4), 366-382.

Gittelman, M. (2006), National institutions, public-private knowledge flows, and innovation performance: A comparative study of the biotechnology industry in the US and France, *Research Policy* 35(7), 1052-1068.

Griliches, Z. (1990), Patent Statistics as Economic Indicators: A Survey, *Journal of Economic Literature* 28(4), 1661-1707.

Griliches, Z. and F. Lichtenberg (1984), Interindustry Technology Flows and Productivity Growth: A Re-examination, *The Review of Economics and Statistics* 66(2), 324-329.

Hagedoorn, J. (1995), Strategic technology partnering during the 1980s: trends, networks and corporate patterns in non-core technologies, *Research Policy* 24(2), 207-231.

Hall, B.H., Jaffe, A.B. and M. Trajtenberg, (2001), *Market Value and Patent Citations: A First Look*, Mimeo, UC Berkeley.

Hall, B.H., Jaffe, A.B. and M. Trajtenberg (2001), The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools, *NBER Working Papers* 8498, National Bureau of Economic Research.

Harhoff, D., Narin, F. Scherer, F.M. and K. Vopel (1999), Citation Frequency and the Value of Patented Inventions, *Review of Economics and Statistics* 81(3), 511-515.

Hu, A.G.Z. and A. B. Jaffe (2003), Patent citations and international knowledge flow: the cases of Korea and Taiwan, *International Journal of Industrial Organization* 21, 849-880.

Jaffe, A.B., Trajtenberg, M. and R. Henderson (1993), Geographic localization of knowledge flows as evidenced by patent citations, *Quarterly Journal of Economics* 108(3), 577-98.

Jaffe, A.B. and M. Trajtenberg (1996), Flows of Knowledge from Universities and Federal Labs: Modeling the Flow of Patent Citations Over Time and Across Institutional and Geographic Boundaries, *NBER Working Papers* 5712, National Bureau of Economic Research.

Jaffe, A.B. and M. Trajtenberg (1999), International Patent Flows: Evidence from Patent Citations, *Economics of Innovation and New Technology* 8, 105-136.

Kaufman, A. and F. Tödling (2001), Science-industry interaction in the process of innovation: the importance of boundary-crossing between systems, *Research Policy* 30(5), 791-804.

Kenney, M. (1986), *Biotechnology: The University-Industrial Complex*, Yale University Press, New Haven, CT.

Ko, Y. (1992), An Economic Analysis of Biotechnology Patent Protection, *Yale Law Journal*. 102, 777-804.

Link, A.N., Siegel, D.S. and B. Bozeman (2007), An empirical analysis of the propensity of academics to engage in informal university technology transfer, *Industrial Corporate Change* 16(4), 641-655.

Lundvall, B.-A. ed. (1992), *National systems of innovation: Towards a theory of innovation and interactive learning*, Pinter, London.

- Maurseth, P.B. and B. Verspagen (2002), Knowledge Spillovers in Europe: A Patent Citations Analysis, *Scandinavian Journal of Economics* 104(4), 531-545.
- Mazzoleni, R. and R.R. Nelson (1998), The benefits and costs of strong patent protection: a contribution to the current debate, *Research Policy* 27(3), 273-284.
- McMillan, G.S., Narin, F. and D.L. Deeds (2000), An analysis of the critical role of public science in innovation: the case of biotechnology, *Research Policy* 29(1), 1-8.
- Merton, R.K. (1973), *The sociology of science: theoretical and empirical investigations*, University of Chicago Press, Chicago.
- Meyer-Krahmer, F. and U. Schmoch (1998), Science-based technologies: university-industry interactions in four fields, *Research Policy* 27(8), 835-851.
- Michel, J. and B. Bettels (2001), Patent citation analysis - A closer look at the basic input data from patent search reports, *Scientometrics* 51(1), 185-201.
- Mowery, D.C., Oxley, J.E. and B.S. Silverman (1996), Strategic alliances and interfirm knowledge transfer, *Strategic Management Journal* 17, 77-91.
- Nelson, R.R. (Ed.) (1993), *National Innovation Systems: A Comparative Analysis*, Oxford University Press, New York.
- Nelson, R.R. and N. Rosenberg (1993), Technical innovation and national systems, in Nelson, R. R. (Ed.), *National Innovation Systems: A Comparative Analysis*, Oxford University Press, Oxford.
- Niosi, J. (2000), Science-based industries: a new Schumpeterian taxonomy, *Technology in Society* 22(4), 429-444.
- Niosi, J. (2003), Alliances are not enough explaining rapid growth in biotechnology firms, *Research Policy* 32(5), 737-750.
- OECD (2006), *OECD Biotechnology Statistics - 2006*, OECD, Paris.
- OECD (2008), *Compendium Of Patent Statistics*, OECD, Paris.

Owen-Smith, J. and W.W. Powell (2001), To Patent or Not: Faculty Decisions and Institutional Success at Technology Transfer, *Journal of Technology Transfer* 26, 99-114.

Pavitt, K. (1984), Sectoral patterns of technical change: Towards a taxonomy and a theory, *Research Policy* 13(6), 343-373.

Pisano, G.P. (1990), The R&D Boundaries of the Firm: An Empirical Analysis, *Administrative Science Quarterly* 35(1) Special Issue: Technology, Organizations, and Innovation, pp. 153-176.

Pyka, A. and P. P. Saviotti (2005), The Evolution of R&D Networking in the Biotechnology-Based Industries, *International Journal of Entrepreneurship and Innovation Management* 5, 49-68.

Porter, M.E. (2000), Location, Competition, and Economic Development: Local Clusters in a Global Economy, *Economic Development Quarterly* 14(1), 15-31.

Porter, M.E. (2000)a, The Economic Performance of Regions, *Regional Studies* 37(6&7), 549-578.

Powell, W.W., Koput, K.W. and L. Smith-Doerr (1996), Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology, *Administrative Science Quarterly* 41(1), 116-45.

Powell, W.W., Koput, K.W., Smith-Doerr, L. and J. Owen-Smith (1999), Network Position and Firm Performance: Organizational Returns to Collaboration, in: Andrews, S. and D. Knoke (eds), *Research in the Sociology of Organizations*, JAI Press, Greenwich, Conn.

Prevezer, M. (2001), Ingredients in the Early Development of the U.S. Biotechnology Industry, *Small Business Economics* 17(1-2), 17-29

Rosenberg, N. (1994), *Exploring the Black Box: Technology, Economics and History*, Cambridge University Press, Cambridge.

Scherer, F.M. (1982), Inter-Industry Technology Flows and Productivity Growth, *The Review of Economics and Statistics* 64(4), 627-634.

Schmookler, J. (1966), *Invention and Economic Growth*, Harvard University Press, Cambridge (Mass.).

Stolpe, M. (2002), Determinants of knowledge diffusion as evidenced in patent data: the case of liquid crystal display technology, *Research Policy* 31(7), 1181-1198.

Thompson, P. and M. Fox-Kean (2005), Patent Citations and the Geography of Knowledge Spillovers: A Reassessment, *The American Economic Review* 95(1), 450-460.

Thompson, P. (2006), Patent Citations and the Geography of Knowledge Spillovers: Evidence from Inventor- and Examiner-added Citations, *The Review of Economics and Statistics* 88(2), 383-388.

Trajtenberg, M., Henderson, R. and A.B. Jaffe (1992), Ivory tower versus corporate lab: An empirical study of basic research and appropriability, *NBER Research Paper* 4146, National Bureau of Economic Research.

Van Zeebroeck, N., van Pottelsberghe de la Potterie, B. and W. Han (2006), Issues in measuring the degree of technological specialisation with patent data, *Scientometrics* 66(3), 481-492.

Von Hippel, E. (1994), “Sticky information” and the locus of problem solving: Implications for innovation, *Management Science* 40(4), 429-439.

Zucker L.G., Darby, M.R. and J. Armstrong (1994), Intellectual Capital and the Firm: The Technology of Geographically Localized Knowledge, *NBER Research Paper* 4653, National Bureau of Economic Research.

Zucker, L.G. and M.R. Darby (1996), Star Scientists and Institutional Transformation: Patterns of Invention and Innovation in the Formation of the Biotechnology Industry, *Proceedings of the National Academy of Sciences of the United States of America* 93(23), 12709-12716.

Zucker, L.G. and M.R. Darby (1998), Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises, *American Economic Review* 88 (1), 290-306.

Appendix

Table A1: T-Tests of the descriptive statistics

Variable	T-test between INDIND and INDSCI mean difference	
<i>TECHCL</i>	0.077	***
<i>CULCL</i>	0.044	***
<i>INTGINI</i>	0.133	*
<i>COMMON_CITED</i>	-0.026	***
<i>ECVALUE</i>	0.553	*
<i>ECVALUESQ</i>	31.135	
<i>NOPATS_CITED</i>	-9.197	**
<i>NOPATS_CITING</i>	46.459	***
<i>YEAR_DIFF</i>	-0.065	
<i>YEAR_DIFFSQ</i>	-1.204	
<i>DE_CITED</i>	0.013	*
<i>US_CITED</i>	-0.018	
<i>JP_CITED</i>	0.031	***
<i>DE_CITING</i>	0.009	
<i>US_CITING</i>	-0.098	***
<i>JP_CITING</i>	0.122	***

Table A2: Marginal effects of the bivariate probit model

Variable	Industry to science (INDSCI)			Industry to industry (INDIND)		
	dy/dx	Std. Err.		dy/dx	Std. Err.	
<i>TECHCL</i> ^a	0.010	0.003	***	0.267	0.009	***
<i>CULCL</i> ^a	0.005	0.004		0.148	0.010	**
<i>INTGINI</i>	-0.003	0.001	**	-0.007	0.003	***
<i>COMMON_CITED</i> ^a	0.088	0.019	***	-0.018	0.026	***
<i>ECVALUE</i>	0.001	0.001		-0.003	0.002	***
<i>ECVALUESQ</i>	0.000	0.000		0.000	0.000	**
<i>NOPATS_CITED</i>	0.000	0.000		0.000	0.000	***
<i>NOPATS_CITING</i>	-0.001	0.000	***	0.000	0.000	**
<i>YEAR_DIFF</i>	-0.001	0.001		0.008	0.003	
<i>YEAR_DIFFSQ</i>	0.000	0.000		0.000	0.000	**
<i>DE_CITED</i> ^a	-0.015	0.004	***	-0.032	0.016	
<i>US_CITED</i> ^a	-0.001	0.004		-0.090	0.012	**
<i>JP_CITED</i> ^a	-0.009	0.004	**	-0.053	0.012	***
<i>DE_CITING</i> ^a	0.016	0.008	**	0.009	0.016	***
<i>US_CITING</i> ^a	0.011	0.004	***	-0.058	0.011	
<i>JP_CITING</i> ^a	-0.024	0.004	***	0.029	0.013	***

Note: ***, **, * indicate a significance level of 1%, 5%, 10%, ^a dummy variable.

Table A3: Correlation matrix

	TECHCL	INTGINI	CULCL	ECVALUE	ECVALUESQ	NOPATS_CITED	NOPATS_CITING	COMMON_CITED	YEAR_DIFF	YEAR_DIFFSQ	DE_CITED	US_CITED	JP_CITED	DE_CITING	US_CITING	JP_CITING
TECHCL	1															
INTGINI	-0.077	1														
CULCL	0.041	0.026	1													
ECVALUE	0.066	0.799	0.042	1												
ECVALUESQ	0.044	0.550	0.042	0.864	1											
NOPATS_CITED	0.012	0.156	0.031	0.153	0.094	1										
NOPATS_CITING	0.016	-0.004	-0.008	0.022	0.047	0.051	1									
COMMON_CITED	-0.017	0.007	0.004	-0.006	-0.013	-0.023	-0.007	1								
YEAR_DIFF	-0.061	-0.040	-0.044	-0.052	-0.053	-0.031	0.093	-0.012	1							
YEAR_DIFFSQ	-0.052	-0.028	-0.035	-0.038	-0.037	-0.032	0.072	-0.011	0.935	1						
DE_CITED	-0.037	-0.066	-0.077	-0.098	-0.059	0.100	0.027	0.086	0.027	0.016	1					
US_CITED	0.060	0.150	0.274	0.179	0.120	0.148	0.005	-0.028	-0.021	-0.014	-0.295	1				
JP_CITED	-0.055	-0.104	0.026	-0.134	-0.078	-0.140	0.001	0.006	-0.011	-0.012	-0.183	-0.465	1			
DE_CITING	0.012	-0.014	-0.111	-0.022	-0.018	0.016	0.280	0.012	0.044	0.031	0.067	-0.025	-0.013	1		
US_CITING	0.003	0.026	0.318	0.049	0.037	0.002	-0.018	-0.006	-0.083	-0.074	-0.023	0.068	-0.057	-0.307	1	
JP_CITING	-0.034	-0.009	0.036	-0.031	-0.023	0.001	-0.099	-0.001	0.073	0.073	-0.010	-0.054	0.125	-0.199	-0,429	1