How Cost, Complexity and Technological Opportunity Affect the Rate of Patenting

Preliminary

Please do not quote without our permission

Georg von Graevenitz, Stefan Wagner, Dietmar Harhoff

February 29, 2008

Abstract

We investigate determinants of patenting, focusing on effects of costs, complexity of technology and technological opportunity. In a theoretical model of patenting it is shown that in complex technologies greater technological opportunity reduces firms' incentives to patent while greater complexity of technology increases patenting incentives. In contrast firms' patenting incentives rise in discrete technologies as technological opportunity increases. Using European patent data a new measure of technological complexity is derived from patent citations. It is shown that patenting conforms to our theoretical model. The theoretical predictions are tested in a panel which allows us to study patenting behaviour of 2074 firms in 30 technology areas over 15 years. Results from GMM estimation indicate that patent thickets exist in 10 of these areas and have important effects on patenting behaviour.

JEL: L13, L49, L63.

Keywords: Patent thickets, Patent portfolio races, Fragmentation .

Acknowledgements: We would like to thank Dirk Czarnitzki, Rene Belderbos, Joachim Winter and Mark Schankerman for comments. Participants at seminars in Leuven, the 2007 Conference on Monte Verita and at the 2^{nd} EPIP Conference in Copenhagen provided valuable feedback on earlier versions of this project. We would like to thank Bronwyn Hall for supplying us with software to consolidate applicant names. The usual disclaimer applies.

1 Introduction

Strong increases in the level of patent applications have been observed at the United States Patent and Trademark Office (USPTO) (Kortum and Lerner (1998) and Hall (2005)) as well as the European Patent Office (EPO) (von Graevenitz et al. (2007)). These "patent explosions" pose serious challenges for existing patent systems and also for competition authorities.[1]

Explanations for the shift in patenting behaviour concentrate on changes in the legal environment, changing management practices, the complexity of important technologies such as semiconductors, greater fecundity of technology and increased strategic behaviour on the part of firms. While it has been shown that most of these factors play a role, there are no formal models of patenting behaviour that explicitly model these influences.[2] This paper provides a model that encompasses complexity and fecundity of technology as well as strategic behaviour. We show the predictions of the model hold using european patent data.

Kortum and Lerner (1998) first investigated the explosion of patenting at the USPTO, which began in 1984 (Hall (2005)). By a process of elimination Kortum and Lerner (1998, 1999) argue that the shift towards increased patenting is mainly the result of changed management practices making R&D more applied and raising the yield of patents from R&D. In contrast, Hall and Ziedonis (2001) who focused on the semiconductor industry argue that the patenting surge is a strategic response to an increased threat of hold-up in complex technologies. This resulted from the "pro-patent" legal environment ushered in after the establishment of the Court of Appeals for the Federal Circuit in the United States (Jaffe (2000)). Both Kortum and Lerner (1998, 1999) and Hall and Ziedonis (2001) find little evidence for the influence of additional technological opportunity as an explanation for increased patenting.

In this paper we develop a model of patenting covering complex and discrete technologies. The model shows how technological opportunity, complexity of a technology and patenting costs jointly determine the rate of patenting. We model the choice between pursuit of new technological opportunities and deepened protection of existing technologies by patenting of "facets" of the technologies. The model shows strategic patenting behaviour implies firms in a complex technology should patent *less* in response to increasing technological opportunity. Additionally, the model indicates that greater technological complexity will raise firms' incentives to patent. These predictions result from strategic interaction of firms using a complex technology: greater technological opportunity reduces the pressure on firms to defend their stake in existing technologies by patenting heavily, whereas greater complexity increases the scope for holdup and raises the need for strategic buildup of patent portfolios.

To test the model we use a comprehensive dataset based on data from the EPO. It com-

[1]For extensive discussions of the policy questions surrounding current functioning of the patent systems in the United States and in Europe refer to National Research Council (2004); F.T.C. (2003); von Graevenitz et al. (2007) and Bessen and Meurer (2008).

[2]Formal models of patenting abound, for a survey of this literature refer to Scotchmer (2005) or Gallini and Scotchmer (2002). Formal models of patenting in patent thickets do not attempt to span both complex and discrete technologies as we do here: Bessen (2004); Clark and Konrad (2005) and Siebert and von Graevenitz (2006). These models usually build on the older patent race literature Lee and Wilde (1980); Reinganum (1989) and Beath et al. (1989).

prises information on patenting behaviour between 1978 and 2003. We construct a measure of blocking in a complex technology based on information specific to European patents. The measure exploits the fact that patent examiners at the EPO indicate which prior patents block or restrict the breadth of a patent application. We count how often three or more firms applied for mutually blocking patents within a three year period. This gives rise to a count of mutually blocking firm *Triples*. The measure allows us to capture effects of complex blocking relationships which can arise in complex technologies even if patent ownership remains relatively concentrated. We find that the measure allows us to identify effects of blocking in a complex technology.

Additionally, a measure of technological opportunity is needed to test our hypotheses. We use the extent to which patents reference non-patent literature for this purpose. (Meyer (2000); Narin and Noma (1985); Narin et al. (1997)) show that the share of references pointing to non-patent literature (mostly scientific publications) can be a good proxy for strength of the science link of a technology. Variation in the strength of the science link within a technology area will indicate how much technological opportunity there is at a given time.

Our paper follows Kortum and Lerner (1998, 1999) and Hall (2005) in considering patenting across the full range of patentable technologies. This allows us to identify differences in patenting behaviour between complex and discrete technologies.

Firms' patenting behaviour is known to be highly persistent, due to the long term nature of firms' R&D investment decisions. We control for the effects of R&D investment decisions on patenting by including a lagged dependent variable in the empirical model. Building on the theoretical model it is shown that this allows us to control for unobserved variation in fixed costs of patenting and in the value of patenting. The model is estimated using systems GMM estimators (Blundell and Bond (1998); Arellano (2003) and Alvarez and Arellano (2003)) to control for endogeneity of the lagged dependent variable as well as our measure of technological opportunity. Evidence from these regressions as well as results from OLS and a fixed effects estimator all support the theoretical predictions we derive from the theoretical model.

Our results complement the descriptive study of patenting at the EPO undertaken by von Graevenitz et al. (2007). They use a set of indicators to identify technology areas in which firms build up patent portfolios for strategic reasons. We extend their work by showing theoretically and empirically how patenting is affected by variation in complexity of a technology and technological opportunity. Additionally, we provide a new measure of blocking complexity. Both studies show strategic patenting behaviour has become very important in technology areas central to productivity growth in recent years (Jorgenson and Wessner (2007)).

Surveys of the use of patents show that traditionally firms mostly protect their innovations through secrecy or lead time (Levin et al. (1987), Arundel and Kabla (1998), Cohen et al. (2000), Arundel (2001), Arundel (2003)). This used to be particularly true for complex product industries. This has changed dramatically in the last decade. Firms in complex product industries such as semiconductors, telecommunications, software and biotechnology have adopted a strategy of building large patent portfolios and invest heavily in amassing patents.[3] Shapiro (2001) investigates the consequences of the ensuing races to build patent portfolios. He argues that firms in some industries are caught in a prisoner's dilemma in which they jointly create "patent thickets" that raise transactions costs and may damage incentives to innovate. These patent thickets arise where many rival firms own patents that must be combined to create individual new products. This gives rise to complex bargaining problems that often cannot be resolved properly.

In this paper we show for the first time to what extent patent thickets also exist within the patent system administered by the European Patent Office (EPO). We find that incidence and complexity of these thickets are increasing. There are important differences between the patent systems administered by the USPTO and the EPO: in particular, it is claimed that examination of patents is more thorough at the EPO and that the opposition system existing there provides a cheaper way for rival firms to weed out weak patents than patent litigation does in the United States (Hall and Harhoff (2004), von Graevenitz et al. (2007)). Therefore, it is not a foregone conclusion that patent thickets also affect the European patent system. Our finding that they do raises important policy questions: what effects are these patent thickets having on competition in the affected sectors? Can the procedures governing patenting in Europe accommodate and regulate strategic patenting behaviour on the scale documented here? Does Competition policy need to take a more active role in regulating use of patents in complex technologies?

Firms' patenting activities and uses of their patent stocks are increasingly the focus of competition policy cases. These cases have arisen in Europe and in the United States and include the dispute between Intel and Intergraph (Shapiro (2003)), a recent case affecting Yamaha and rivals in the personal watercraft industry (Rubinfeld and Maness (2005)) and last but not least the ongoing dispute between Qualcomm and Nokia. These legal disputes have all revolved around attempts of a dominant patent owner to extract licensing royalties from rivals through aggressive assertion of their patent portfolio. Two of these cases originate in the semiconductor and telecommunications industries. This is no accident: the complexity of the technology employed in this field combined with a very high rate of patenting activity leads to a dispersal of patent rights among rival firms.

This paper is structured as follows. Section 2 provides a theoretical model of patenting which explains how firms' patenting strategies evolve in response to increased patenting by their rivals. We derive three hypotheses from this model that are empirically testable. In section 3 we describe our dataset and the variables we employ to analyse firms' patenting behaviour. As there is little cross industry evidence of patenting trends at the EPO, section 3 also provides a descriptive analysis of these trends, focusing particularly on our measure of complexity and alternative measures thereof. Section 4 provides empirical results and Section 5 concludes.

[3] In the case of semiconductors the reasons for this change are discussed by Grindley and Teece (1997), Jaffe (2000), Hall and Ziedonis (2001)

2 Incentives to patent in discrete and complex technologies

In this section we model firms' patenting behaviour. In particular, we analyse how firms' profit maximising patenting decisions are influenced by the cost of patenting, existing technological opportunity and the complexity of the technology area in which firms patent. Before presenting our formal model we briefly describe the mechanisms modelled below.

Previous literature has shown that firms' efforts to accumulate patents are strongly influenced by characteristics of a technology area such as fragmentation of ownership rights or potential threat of being held up by other patentees. It has been argued that these effects led to a surge in patenting in complex technologies areas such as semiconductors or telecommunications (Hall and Ziedonis (2001), Ziedonis (2004)). These studies, however, do not provide an explicit model of the interaction of complexity and firms' patenting efforts. We model firms' patenting efforts as a function of the complexity of the underlying technology. In order to do so, we propose a simple model of complexity based on the widespread notion that in complex technologies products relate to a (potentially large) number of patents held by various different patentees whereas in discrete technologies a direct product-patent link dominates. In order to create this measure of complexity, we distinguish technological opportunities O representing separate subtechnologies within a technology area. For example, a technological opportunity might be constituted by research related to the development of a certain chemical compound in organic chemistry, the search for a drug in the pharmaceutical area or the development of special circuit in electronics. Complexity within these technological opportunities can arise if it is possible to patent different facets F within an opportunity. If only one facet of an opportunity can be patented, the technology is discrete. At least two facets must be patentable in order to allow for situations where different patentees own patent rights related to the same technology – this is our definition of complexity. An increase in the number of patentable facets increases the potential number of patentees owning patents relating to the same technological opportunity. Hence, we model complexity of a technology as the number of patentable facets. Figure 1 presents a graphical representation of this idea.

Further, we assume that patenting allows firms to benefit from the total value (V) of a technology opportunity. To capture maximum value from the technology opportunity a firm must obtain as many patents as possible on facets of the opportunity. Firms face a tradeoff between patenting more facets per opportunity and patenting more different technological opportunities. We show that technological opportunities and facets are complements in firms' investment decisions.

As the number of facets per opportunity grows, so does the probability that different firms will own patents related to the same opportunity. These firms may need to disentangle their ownership rights, giving rise to legal costs (L). We do not explicitly model the bargaining process between firms that own patents on the same technological opportunity. The literature on patent thickets and complex technology shows that there are many institutional arrangements that allow firms to disentangle overlapping property rights - these include licensing, patent pools, standard setting as well as litigation (Shapiro (2001)). We adopt a reduced form rep-

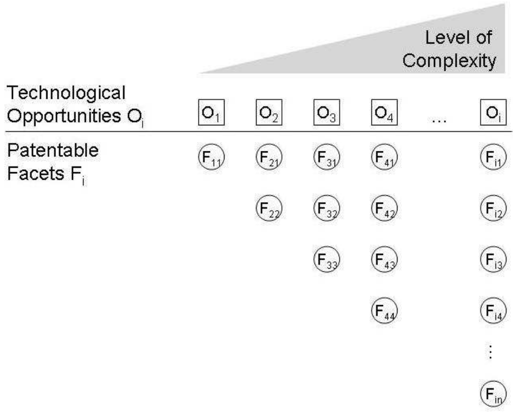

Figure 1: Relation between complexity and the number of patentable facets per technological opportunity. Note that O_1 is discrete by definition as there is no chance of overlapping ownership rights in this technology.

resentation of all of these mechanisms, by assuming that they become more expensive as the number of parties involved increases. Additionally, we assume that firms who own a greater share of patents on a technological opportunity have advantages in bargaining which allow them to appropriate a larger share of the benefits associated with the opportunity. These assumptions are consistent with the arguments advanced by Ziedonis (2004) to explain patent portfolio races in the semiconductor industry.

We assume that the total set of patentable facets in a technology (Ω) consists of O technology opportunities and F facets such that: $FO = \Omega$. The patent office grant only one patent per facet. Each firm knows that there is a contest for patents on the facets of a technological opportunity. This implies that the probability of obtaining a patent is inversely proportional to the number of rivals seeking a patent on the same facet. Firms simultaneously determine the number of opportunities $O_i \in [0, O]$ to invest in and the number of facets $F_i \in [0, F]$ per opportunity which they seek to patent.

Costs and benefits of patenting

Three types of cost are associated with patenting in our model:

i For each opportunity a firm invests in, it faces a fixed cost of R&D: C_o.

ii For each facet which a firm patents the firm faces costs of administering and enforcing the patent if it is granted: C_a.

iii The coordination of R&D on different technologies imposes costs $C_c(O_i)$. We assume that $\frac{\partial C_c}{\partial O_i} > 0$.

The benefits of patenting are a function of the value of each technological opportunity V and the expected number of facets f_i each firm receives a patent on. Define the expected share of facets per patent which each firm obtains as $f_i \equiv \frac{F_i p}{F}$ where F_i is the number of facets each firm invests in per opportunity and p is the probability of winning a patent on a given facet. Note that $f_i \in [0, 1]$. The probability of obtaining granted patent on a given facet is:

$$p = \frac{1}{1 + \frac{\sum_{j \neq i} F_j O_j}{FO}} \quad . \tag{1}$$

This definition of the probability of obtaining a patent on a facet of a technology opportunity reflects our assumption that there is a contest between several firms for each such patent. Then the probability of obtaining the patent depends on the number (n) of rival firms simultaneously trying to obtain the patent. Each firm vying for a patent on a facet will win that patent with $p = \frac{1}{1+n}$. In the expression above we assume that all rival firms make $\sum_{j \neq i} F_j O_j$ patent applications. Dividing these by the set of all patentable facets FO we obtain the number of rivals' patent applications that compete with each firm's own applications.

Given these costs and benefits the expected value of patenting in a technology area is:

$$\pi_i = O_i \Big[V\omega(f_i) - L(f_i, N) \Big] - O_i c_o - O_i F_i p C_a - C_c(O_i) \quad , \tag{2}$$

where total legal costs of owning patents on an opportunity are $L(f_i)$ which decrease in the share of facets owned on that opportunity. $\omega(f_i)$ represents the share of value of a technological opportunity obtained by firm i. It is an increasing function of the firm's share of patents held on a given opportunity.

Comparative statics of this model

To simplify the derivation of comparative statics results we show that the game firms are playing is supermodular. Then we use results on supermodular games to derive comparative statics results [Milgrom and Roberts (1990), Vives (1990, 1999)].[4] We define a symmetric game in which firms' payoffs depend on own strategies and the aggregate strategy of their rivals. Additionally we will assume that strategy spaces are compact. These assumptions imply that only symmetric equilibria exist (Vives (1999)). Additionally, we can characterize the comparativs statics for these equilibria by considering cross-partial derivatives.

We begin by characterising the game firms are playing:

- There are N firms.
- Each firm chooses the number of technological opportunities $O_i \in [0, O]$ and facets $F_i \in [0, F]$ to invest in. The firms' strategy sets S_n are elements of R^2.
- Each firm has the payoff function π_i, defined in equation (14), which is twice continuously differentiable and depends only on rivals' aggregate strategies.

Firms' payoffs depend on their rivals' aggregate strategies because the probability of obtaining a patent on a given facet is a function of the sum of rivals' patent applications $\sum_{i \neq j} F_j O_j$.

[4]For additional expositions of this method refer to Carter (2001) or Amir (2005).

We can show that:

Proposition 1

The game is a smooth supermodular game.

To prove this proposition we must show that the firms' profit functions are supermodular (i) in their own actions and (ii) in every combination of their own actions with those of rival firms [Milgrom and Roberts (1990)].

To begin with we derive the first order conditions characterising the optimal number of technological opportunities and facets firms invest in:

$$\frac{\partial \pi}{\partial O_i} = V\omega(f_i) - L(f_i) - C_o - F_i p C_a - \frac{\partial C_c}{\partial O_i} = 0 \tag{3}$$
$$\frac{\partial \pi}{\partial F_i} = \Big[V\frac{\partial \omega}{\partial f_i} - \frac{\partial L}{\partial f_i} - FC_a\Big] O_i \frac{p}{F} = 0 \tag{4}$$

These first order conditions constitute a system of implicit relations which determine the optimal choice of opportunities ($\hat{O}_i$) and facets ($\hat{F}_i$) chosen by each firm in equilibrium.

Given this system of first order conditions we can show that firms' profit functions are supermodular. To see this we derive the cross partial derivatives with respect to firms' own actions as well as those of rival firms:

$$\frac{\partial^2 \pi_i}{\partial O_i \partial F_i} = V\frac{\partial \omega}{\partial f_i}\frac{p}{F} - \frac{\partial L}{\partial f_i}\frac{p}{F} - pC_a = 0 \tag{5}$$

Notice that this expression must be zero as it can be transformed to the first order condition (4) for the optimal number of facets by multiplication with O_i. Next consider effects of rivals' actions on firms' own actions:

$$\frac{\partial^2 \pi_i}{\partial O_i \partial O_j} = V\frac{\partial \omega(f_i)}{\partial f_i}\frac{F_i}{F}\frac{\partial p}{\partial O_j} - \frac{\partial L(f_i)}{\partial f_i}\frac{F_i}{F}\frac{\partial p}{\partial O_j} - F_i C_a \frac{\partial p}{\partial O_j} = 0 \quad , \tag{6}$$
$$\frac{\partial^2 \pi_i}{\partial O_i \partial F_j} = V\frac{\partial \omega(f_i)}{\partial f_i}\frac{F_i}{F}\frac{\partial p}{\partial F_j} - \frac{\partial L(f_i)}{\partial f_i}\frac{F_i}{F}\frac{\partial p}{\partial F_j} - F_i C_a \frac{\partial p}{\partial F_j} = 0 \quad , \tag{7}$$
$$\frac{\partial^2 \pi_i}{\partial F_i \partial O_j} = \Big[V\frac{\partial \omega}{\partial f_i} - O_i\frac{\partial L}{\partial f_i} - FC_a\Big]\frac{O_i}{F}\frac{\partial p}{\partial O_j} + \Big[O_i V\frac{\partial^2 \omega}{\partial {f_i}^2} - O_i\frac{\partial^2 L}{\partial {f_i}^2}\Big]\frac{pF_i}{F^2}\frac{\partial p}{\partial O_j} > 0 \quad , \tag{8}$$
$$\frac{\partial^2 \pi_i}{\partial F_i \partial F_j} = \Big[V\frac{\partial \omega}{\partial f_i} - \frac{\partial L}{\partial f_i} - FC_a\Big]\frac{O_i}{F}\frac{\partial p}{\partial F_j} + \Big[O_i V\frac{\partial^2 \omega}{\partial {f_i}^2} - O_i\frac{\partial^2 L}{\partial {f_i}^2}\Big]\frac{pF_i}{F^2}\frac{\partial p}{\partial F_j} > 0 \quad , \tag{9}$$

where the first two conditions are transformations of the first order condition for the optimal number of facets (4). In case of the lower two conditions notice that the first term in square brackets is zero as it is just that same first order condition. The terms in the second set of brackets are negative if:

i) the share of value of a technological opportunity which a firm can appropriate with additional facets is decreasing as firms' share of facets on a technological opportunity increases: $\frac{\partial^2 \omega}{\partial {f_i}^2} \leq 0$;

ii) legal costs fall at a decreasing rate as firms' share of facets on a technological opportunity increases: $\frac{\partial^2 L}{\partial {f_i}^2} \geq 0$.

At least one of these two conditions must be fulfilled for the game outlined above to be smooth supermodular.

Condition (i) indicates that as a firm's share of patents on a technological opportunity increases, the marginal value of additional patents is decreasing. This assumption will hold if a firm holding some patents on a technological opportunity is able to make use of the technology covered to some extent in the face of blocking patents.[5]. In contrast if any one patent on a technological opportunity blocks the use of the technology entirely, the assumption is violated.[6]

Condition (ii) indicates that firms' legal costs of appropriating a share of the value of a technology opportunity fall if they own a larger share of patents on that technology opportunity. This assumption reflects the widespread belief that larger patent portfolios are beneficial to firms operating in technology areas that fall within complex technologies because they provide firms with bargaining chips (Hall and Ziedonis (2001)).

Note that the game will not be smooth supermodular if the technology is not complex. By definition in that case there is only one facet $(F = 1)$ per technological opportunity. Then firms appropriate the whole value of the technological opportunity with one patent and the second derivatives in (8) and (9) are zero. We will return to this case below.

Now we turn to the comparative statics effects of an increase in technological opportunity on firms' actions. We show that:

Proposition 2

Increased technological opportunity reduces firms' patenting efforts in a complex technology.

To determine the effects of an increase in technological opportunity O we investigate the following cross-partial derivatives:

$$\frac{\partial^2 \pi_i}{\partial O_i \partial O} = \left[V\frac{\partial w}{\partial f_i} - \frac{\partial L}{\partial f_i} - FC_a\right]\frac{\partial p}{\partial O}\frac{F_i}{F} = 0 \tag{10}$$

$$\frac{\partial^2 \pi_i}{\partial F_i \partial O} = \left[V\frac{\partial \omega}{\partial f_i} - \frac{\partial L}{\partial f_i} - FC_a\right]\frac{O_i}{F}\frac{\partial p}{\partial O} + \left(O_i V\frac{\partial^2 \omega}{\partial {f_i}^2} - O_i\frac{\partial^2 L}{\partial {f_i}^2}\right)\frac{pF_i}{F^2}\frac{\partial p}{\partial O} < 0 \tag{11}$$

The terms in square brackets in both expressions above are zero by the first order condition (4) for the optimal number of facets. The term in round brackets in equation (11) is negative if the game is smooth supermodular, i.e. if the technology is complex.

Therefore, greater technological opportunity lowers firms' overall investments in patenting. It reduces the intensity of competition to dominate individual technological opportunities which lowers investments in facets and the number of new technologies which firms invest in.

Now we turn to the question how an increase in the complexity of a technology affects firms' incentives to patent. We find that the effect is ambiguous and depends on the relative

[5]Such a setting is modelled in Siebert and von Graevenitz (2008, 2006)

[6]Clark and Konrad (2005) make such an assumption.

strength of two effects: the costs of administering more patents and the marginal benefits of additional patents. Only if these marginal benefits are high enough will the term be positive.

To see this consider the following cross-partial derivatives:

$$\frac{\partial^2 \pi_i}{\partial O_i \partial F} = \left[V\frac{\partial w}{\partial f_i} - \frac{\partial L}{\partial f_i} - FC_a\right]\frac{\partial p}{\partial O}\frac{\partial f_i}{\partial F} = 0 \tag{12}$$

$$\frac{\partial^2 \pi_i}{\partial F_i \partial F} = \left[V\frac{\partial \omega}{\partial f_i} - \frac{\partial L}{\partial f_i} - FC_a\right]\frac{O_i p^2}{FO} + \left(V\frac{\partial^2 \omega}{\partial {f_i}^2}\frac{\partial f_i}{\partial F} - \frac{\partial^2 L}{\partial {f_i}^2}\frac{\partial f_i}{\partial F} - C_a\right)\frac{O_i}{F_i}f_i \tag{13}$$

Here the terms in square brackets are zero by the first order condition (4) for the optimal number of facets. The term in round brackets in equation (13) is positive if the costs of administration of patents C_a are insignificant.

This shows that:

Proposition 3

Increased complexity of a technology will increase firms' patenting efforts if the costs of administering patents are low relative to their value as bargaining chips.

Finally, consider again the case of a discrete technology opportunity. Here $F = F_i = 1$ by definition. Therefore firms's payoffs are defined as:

$$\pi_i = O_i V p - O_i c_o - O_i p C_a - C_c(O_i) \quad . \tag{14}$$

We have already noted that a game with this payoff function is no longer supermodular. However we can show that under the slightly stronger assumption that costs of coordinating technological opportunities ($C_c(O_i)$) are strictly convex in the number of opportunities firms invest in, we obtain a unique equilibrium for the game. We can then demonstrate that:

Proposition 4

Greater technological opportunity increases firms' patenting efforts in a discrete technology.

To see that this is true consider the first and second order derivatives of the payoff function with respect to technological opportunities invested in:

$$\frac{\partial \pi}{\partial O_i} = (V - C_a)p - \frac{\partial C_c}{\partial O_i} = 0 \qquad \frac{\partial^2 \pi}{\partial {O_i}^2} = -\frac{\partial^2 C_c}{\partial {O_i}^2} \quad . \tag{15}$$

If we assume that costs of coordianting technological opportunities are strictly convex: $\frac{\partial^2 C_c}{\partial {O_i}^2} > 0$, then Proposition 4 can be proved with the help of the implicit function theorem:

$$\frac{\partial O_i}{\partial O} = -\frac{\partial^2 \pi}{\partial O_i \partial O} \Big/ \frac{\partial^2 \pi}{\partial {O_i}^2} > 0 \quad , \tag{16}$$

where $\frac{\partial^2 \pi}{\partial O_i \partial O} = (V - C_a)\frac{\partial p}{\partial O} > 0$.

This concludes our analysis of the model.

3 Description of the dataset and of important patenting trends

In this section we discuss the data used to test our theoretical model. In particular, a new measure of complexity of a technology is discussed. Next, we provide descriptive evidence supporting the theoretical model. Discrete and complex technology areas are compared with regard to selected patent indicators. Using our measure of complexity we show that descriptive evidence on patenting provides support for the theoretical model.

3.1 Dataset and derivation of variables

Our empirical analysis is based on the PATSTAT database ("EPO Worldwide Patent Statistical Database") provided by the EPO.[7] This database includes data on about 56 million patent applications filed at more than 65 patent offices world-wide. It contains procedural and bibliographic information on patents including information on referenced documents (patent citations). We analyse all patent applications filed at the EPO between 1980 and 2003 – more than 1,5 million patent applications with about 4.5 million referenced documents.

We classify patents using the IPC classification which allows us to analyse sectoral differences in patenting activities. The categorisation used is based on an updated version of the OST-INPI/FhG-ISI technology nomenclature.[8] This classification divides the domain of patentable technologies into 30 distinct technology areas.[9] We also classify selected technology areas as discrete or complex using to the classification of Cohen et al. (2000). This classification received additional support in Hall (2005).

Below we show that there are clear differences between complex and discrete technologies on the basis of this distinction. However, we also provide a new continuous variable that captures the degree of complexity of technologies. We show that there are some differences between this variable and the classification suggested by Cohen et al. (2000).

In the following we discuss our measures of patenting, technological opportunity and complexity. These are the most important variables needed to test the theoretical model. Additionally, we discuss several variables that will be used as control variables in the empirical model of section 4. These describe additional influences on firms' patenting intensity.

Measures of patenting, complexity and technological opportunity

Number of patent applications We compute the number of patent applications A_{iat} filed by applicant i separately for all OST-INPI/FhG-ISI 30 technology areas a on an annual (t) basis. To aggregate patent applications to the firm level two challenges must be overcome: firm names provided in PATSTAT are occasionally misspelled and subsidiaries of larger firms are not identified in the dataset. Therefore, we devoted a considerable amount of resources to

[7] We currently use the September 2006 version of PATSTAT.

[8] See OECD (1994), p. 77

[9] These are listed in Table 6 in the appendix

clean applicant names and to consolidate ownership structures.[10] The aggregation of patent applications are based on these consolidated applicants' identities. The variables discussed below are also based on this consolidation.

Due to the skew distribution of patent applications we transform the variable logarithmically to derive a dependent variable for estimation. Table 2 shows the transformed variable is much closer to a normally distributed variable than the raw measure of patent applications.

Technological opportunity In our model, we establish a clear relationship between firms' patenting levels in complex technologies and the emergence of new technological opportunities. Unfortunately, a direct measure of existence or emergence of new technological opportunities does not exist. Instead, we use a construct that is based on the strength of the link between R&D firms conduct within a technology area and relevant basic research as an indirect measure of the emergence of new technological opportunities. This construct is based on the assumption that basic research is more likely to open up new technological opportunities than applied research which predominantly refines existing technologies.

Early stages of the evolution of a technology are characterised by a large share of basic research often conducted in publicly-funded labs. In later stages of a technology industry driven development of existing technological opportunities will dominate basic research. Then the focus is on refining existing opportunities rather than creating new ones. While there is no perfect measure for the position of a technology area in the stylised cycle of technology evolution, the share of references listed on a patent which point to non-patent literature (mostly scientific publications) can be used as a good proxy for the strength of the science link of a technology (Meyer (2000); Narin and Noma (1985); Narin et al. (1997)).

Therefore, we use the share of non-patent references relative to all references contained on a patent as a proxy for a patent's position in the technology cycle and hence as a measure for the creation of new technological opportunities. As we are interested in the characterisation of technological areas with regard to the existence of new technological opportunities, we compute the average of the share of non-patent references relative to all references on a patent on the level of OST-INPI/FhG-ISI area a and year t for our multivariate analyses.

Complexity of technology areas The distinction between discrete and complex technologies is widely accepted in the literature (Cohen et al. (2000), Kusonaki et al. (1998), Hall (2005)). Discrete technologies are characterised by a relatively strong product-patent link, e.g. in pharmaceuticals or chemistry, whereas in complex industries products are likely to build upon technologies protected by a large number of patents held by various parties. It is often held that patent filing strategies vary largely between discrete and complex industries.

[10]The aggregation of patenting activities on the firm levels involved great efforts consolidating subsidiaries of large corporations. Detailed information on the cleaning and aggregation algorithms can be obtained from the authors upon request. We would like to thank Bronwyn Hall for providing us with software for this purpose. We used this and undertook additional efforts to consolidate firm names.

Despite the widely acknowledged notion of a technology's complexity there is no direct measure of it nor is there an indirect construct related to complexity. Kusonaki et al. (1998) and Cohen et al. (2000) (footnote 44) provide schemes which classify industries as discrete or complex based on ISIC codes. These classification schemes are based on qualitative evidence gathered by the authors from various sources in order to separate different industrial sectors into complex or discrete areas. A major drawback of a classification based on prior information from industry codes is that is does not allow to analyse the influence of different levels of complexity but only to distinguish the binary cases discrete and complex.

To improve on this, we measure complexity of a technology area through firms' patenting activities. Our measure is derived from to the degree of overlap between firms' patent portfolios. Such overlap leads to blocking dependencies among firms. If existing patents containing prior art critical to the patentability of new inventions in a field are held by both firms, each firm can block its rival's use of innovations. Then, a firm can only commercialise a technology if it gets access to a rival's patented technology. In areas where products draw on technological opportunities protected by numerous firms (complex technologies) we expect to observe a large number of such dependencies. In discrete technologies the inverse should be true.

We capture blocking dependencies among firms by analysing the references contained in patent documents. References to older patents or to non-patent literature are included in EPO patents in order to document the extent to which inventions satisfy the criteria of patentability (Harhoff et al. (2006)). Often, existing prior art limits patentability of an invention. For example, the existence of an older but similar invention can reduce the patentability of a newer invention. In these cases *critical* documents containing conflicting prior art are referenced in patent documents and are classified as X or Y references by the patent examiner at the EPO during the examination of the patent application.[11] If the patentability of a firm A's inventions is frequently limited by existing patents of another firm B, it is reasonable to assume that the R&D of A is blocked by B to a certain degree. If the inverse is also true, A and B are in a mutual blocking relationship which we call a blocking pair. If more than two firms own mutually blocking patents the complexity of blcoking relationships increases and resolution of blocking becomes increasingly costly. To capture more complex structures of blocking we compute the number *Triples* in which three firms mutually block each other's patents. Figure 2 provides a graphical example of our complexity measure.

From a computational perspective, pairs and triples are identified using the following approach: For each firm i we analyse all critical patent references contained in firm i's patents applied for in a technology area a over the current and the two preceding years ($t-2$ to t) and identify the owners of the referenced patent documents. In the next step we keep the most frequently referenced firms (top 20) yielding annual lists of firms which are blocking firm i in

[11] A patent contains various different types of references – not all of them are critical. Often, related inventions which are not critical for the patentability of the invention seeking patent protection are also included in the patent document. The EPO provides a full classification of the references included in patent documents allowing us to identify critical references which are classified as X or Y. Table A contains an overview of different types of references contained in patents.

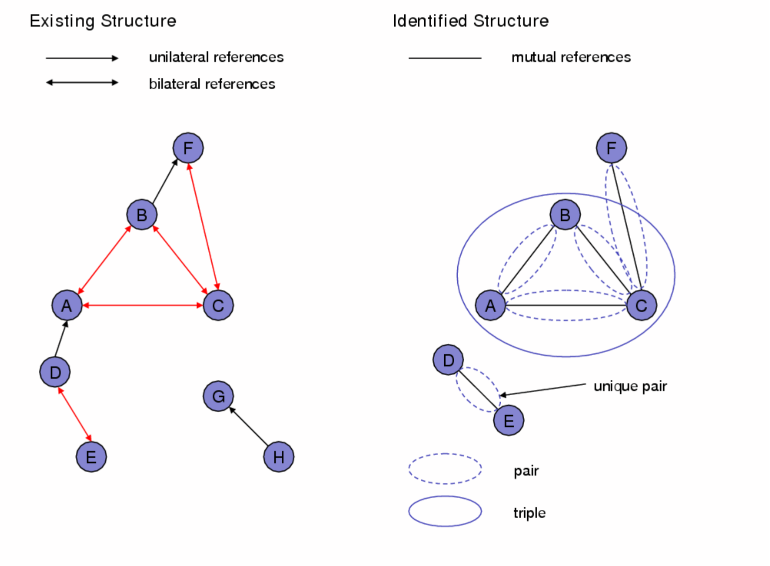

Figure 2: Identification of our measures of a technology field's complexity.

year t.[12] Pairs are then established if firm A is on firm B's list of most frequently referenced firms and, at the same time, firm B is on firm A's list of most frequently referenced firms. Finally, triples are formed if firm A and firm B, firm A and firm C and firm B and firm C form pairs in the same year. We include the total number of existing triples$_{at}$ in area a and year t in our regression in order to analyse how the complexity of a technology area influences firms patenting behaviour in this area.

Control variables

Fragmentation of prior art Ziedonis (2004) showed that semiconductor firms increase their patenting activities in situations where patent holdings are largely fragmented across different parties. Ziedonis' fragmentation index has predominantly been studied in complex industries (Ziedonis (2004), Schankerman and Noel (2006)) where increasing fragmentation has been found to increase the number of firms' patent applications. This has been attributed to firms' efforts to reduce potential hold-up by opportunistic patentees owning critical or blocking patent rights – a situation which is often associated with the existence of *patent thickets*.

We construct an index of fragmentation of patent ownership for each firm based on the fragmentation index proposed by Ziedonis (2004):

$$Frag_{iat} = 1 - \sum_{j=1}^{n} s_{ijt} \tag{17}$$

where s_{ijt} is firm i's share of critical references pointing to patents held by firm j. Small

[12] The threshold of keeping only the 20 most frequently referenced patent owners is an arbitrary choice. Our results are robust to different choices of the threshold level.

values of this fragmentation index indicate that prior art referenced in a firm's patent portfolio is concentrated among few rival firms and vice versa.

Unlike previous studies of patenting in complex technologies relying on USPTO patent data (Ziedonis (2004),Schankerman and Noel (2006)) we base the computation of the fragmentation index solely on critical references which are classified as limiting the patentability of the invention to be patented (X and Y references). This distinction is not available in the USPTO data. Computing the fragmentation index based on critical references should yield a more precise measure of the hold up potential associated with fragmentation of patent holdings in a technology area.

Technological diversity of R&D activities A firm's reaction to changing technological or competitive characteristics in a given technology area might be influenced by its opportunities to strengthen its R&D activities in other fields. For example, if a firm is active in two technology areas it might react by a concentration of its activities in one area if competition in the other area is increasing. If a firm is active in only one technology area, it does not similar possibilities to react to increases in competitive pressure. In order to control for potential effects of opportunities to shift R&D resources we include the total number of technology areas ($\text{Areas}_{i,t}$) with at least one patent application filed by firm i in year t.

Size dummies. While we do not explicitly model the influence of firm size on patenting behaviour, it seems reasonable to assume that the cost of obtaining and upholding a patent depends on the size of a firm. In particular, larger firms might face lower legal cost due to economies of scale, increased potential to source in legal services and accumulation of relevant knowledge which in turn might lead to a different patenting behaviour than smaller firms. For instance Somaya et al. (2007), find that the size of internal patent departments positively influences firms' patenting propensity.

If the economies-of-scale argument holds, the cost of patenting should not be directly related to size characteristics such as a firm's number of employees, its total revenues or sales. Rather, the cost of patenting can be assumed to be a function of the total amount of patents filed by a firm. Therefore, we include a 'size dummy' variable based on the number of patents filed by a firm in a technology area in a given year in our regressions. We distinguish between small and large patentees. These size categories are based on annual patent applications in a given area a. Firms belonging to the upper half of the distribution of patentees in a given year are coded as large firms.

3.2 Descriptive analysis of patenting in Europe

In this section we provide descriptive aggregate statistics on patenting trends at the EPO. We establish several stylized facts which provide support for our modelling approach.

Figure 3 presents annual patent applications filed at the EPO between 1978 and 2003. We distinguish applications filed in complex and discrete technology areas using the categorisation

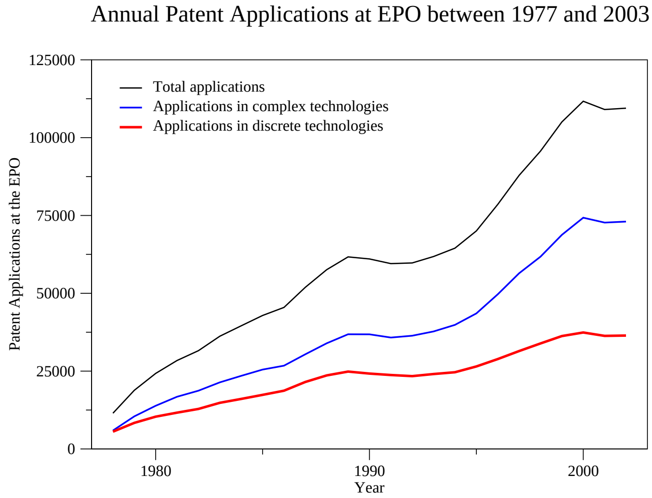

Figure 3: Annual number of patent applications filed at the EPO by priority year. Note: Blue line indicates total patent applications. Red line indicates patent applications in complex technology areas. Green line indicates patent applications in discrete technology areas.

of Cohen et al. (2000). The Figure shows patenting grew strongly over this period, with the main contribution coming from technology areas classified as complex. This development is comparable to trends at the USPTO. Hall (2005) shows that the strong increase in patent applications is is driven by firms patenting in the electrical, computing and instruments area all of which are complex technology areas by the classification of Cohen et al. (2000).

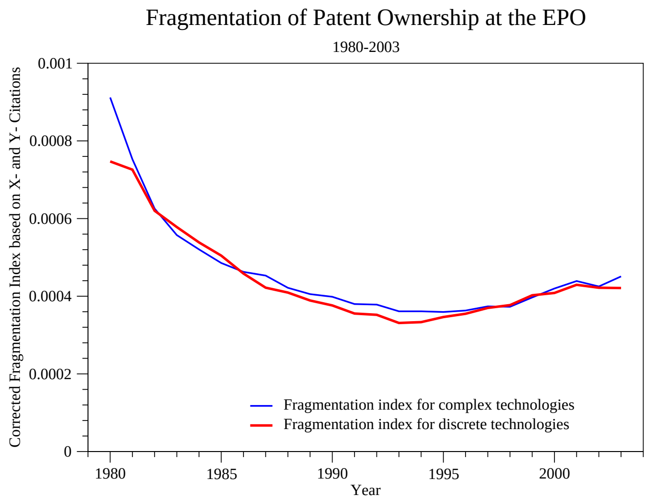

Figure 4: Average fragmentation index. Note: Blue line indicates average level of fragmentation index in complex technology areas. Red line indicates average level of fragmentation index in discrete technology areas.

Now we turn to explanations for the strong growth in patenting. First, consider a leading explanation for increased patenting in complex technology areas: the fragmentation of patent rights in a complex technology area is likely to raise firms' transactions costs as they must bargain with increasing numbers of rivals in order to prevent hold up of their products. Ziedonis (2004) and Schankerman and Noel (2006) show that increased fragmentation of patents leads to greater patenting efforts in the semiconductor and software industries respectively. Figure 4 provides annual averages of the fragmentation index at the EPO for the years 1980 to 2003.[13] Two observations derived from Figure 4 are striking: First, fragmentation of ownership rights fell steadily before 1995 and then increased gradually thereafter. Second, the difference in the fragmentation index in complex and discrete technology areas is negligible.

Both observations raise the question whether the growth in patent applications can be attributed to fragmentation alone. While the development of fragmentation in complex and discrete areas is almost identical we observe striking differences in the growth of patent applications between complex and discrete technology areas.

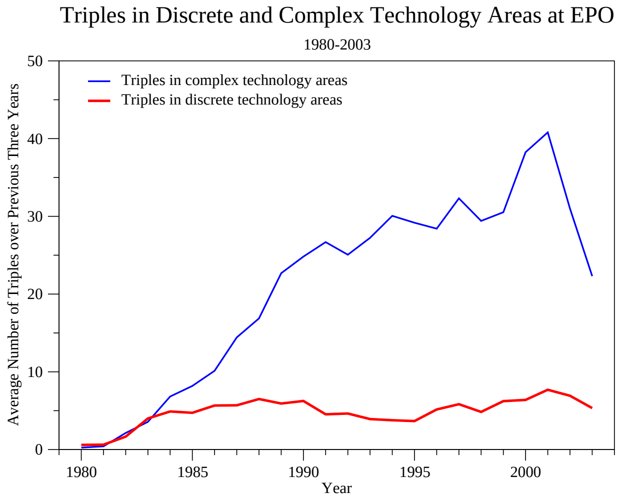

Figure 5: Average number of triples identified. Note: The blue line indicates average number of triples in complex technology areas. The red line indicates average number of triples in discrete technology areas.

Next we explore two explanations for the increase in patenting at the EPO that build on the theoretical model developed above: firstly firms build patent portfolios to strengthen their bargaining positions if complex bargaining situations are more likely to arise and secondly the pressure to obtain patents becomes more intense as technological opportunity declines. The first of these explanations is similar to the explanation for patenting derived from fragmentation of property rights: it also emphasises transactions costs increases derived from bargaining over blocking patents. However, we believe that transactions costs also rise if a small number of firms own patent rights that depend on the rights of other firms that also block each other. Then, bargaining will become increasingly complex as blocking cannot be resolved through a series of bilateral negotiations. Our measure of mutual blocking between three and more firms (Triples) captures the degree to which complex blocking arises.

In Figure 6 this measure is presented. The Figure presents annual averages of the number of Triples in complex and in discrete areas.[14] We observe very different developments of the count of Triples in these two fields. The number of Triples remains largely stable at values

[13]The precise definition of this measure is given in Section 3.1 above.

[14]We distinguish complex and discrete using the classification suggested by Cohen et al. (2000) here.

well under 10 in discrete technology areas, while it increases strongly in complex technology areas. It is reassuring to see that our measure of complex bargaining situations is greater in complex technologies as previously defined by Cohen et al. (2000).

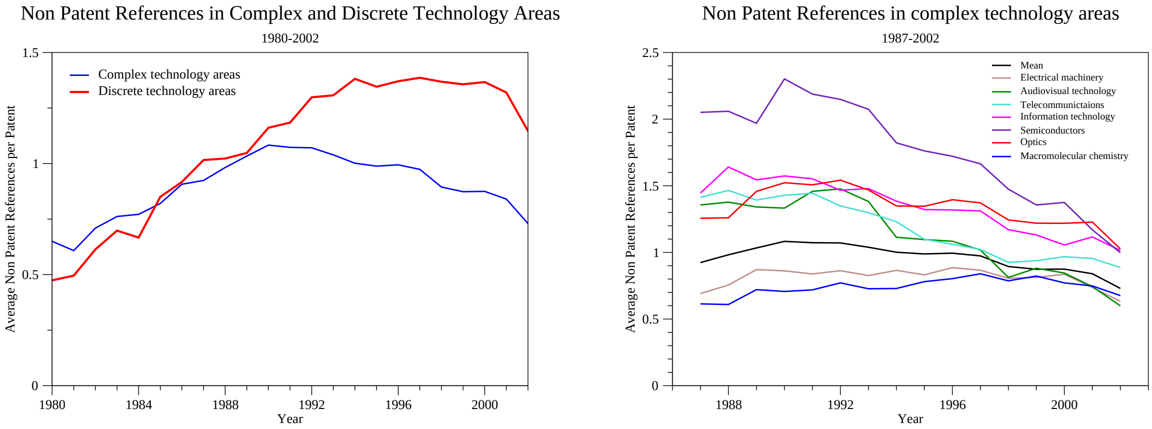

Figure 6: The left panel presents average non patent references per patent for complex (blue line) and discrete (red line) technology areas. The right panel presents average non patent references per patent for several complex technology areas. This panel focuses only on the sample period we use for our regressions below.

This shows that blocking intensities almost certainly contributed to the strong increases in patenting that we observe in Figure 3. Next we turn to the development of technological opportunity. In our theoretical model Proposition 2 indicates greater technological opportunity in a complex technology should lower the pressure to patent. As noted in Section 3.1 we measure technological opportunity using changes in the rate of references to non patent literature within a technology area. This measure will provide information about variation in technological opportunity between and across technology areas. The left panel of Figure 3.2 shows that technological opportunity was generally greater in discrete technology areas after 1990, than in complex technology areas. The right hand panel of the Figure shows that the average level of non patent references in complex technology areas masks considerable variation across and especially within complex technologies over time.

Note that the level of non patent references in the complex technology areas began to decline just after 1992, which coincides with the date at which the growth in patent applications at the EPO picked up as Figure 3 shows. These descriptive results suggest that a multivariate analysis of patenting levels based on the theoretical model presented above will prove to be interesting.

To complete this section Table 1 provides additional information on the distribution of Triples across all 30 technology areas. This shows how significant the hold up potential measured by Triples is in the ICT technologies. The number of Triples is between five and six times as large there as it is in other industries such as Handling, Printing which still exhibit significant complexity.

Table 1: The Distribution of Triples Between 1987 and 2002

Technology area	**Mean**	**Median**	**Std. dev.**	**Minimum**	**Maximum**
Electrical machinery, Electrical energy	24.23	20	8.99	10	42
Audiovisual technology	116.48	120	17.68	74	148
Telecommunications	99.64	93	39.17	27	166
Information technology	57.16	59	10.71	28	73
Semiconductors	62.84	63	17.89	26	91
Optics	57.30	58	12.02	42	77
Analysis, Measurement, Control	6.61	4	6.31	0	21
Medical technology	4.10	3	2.16	1	8
Nuclear engineering	0.95	1	1.17	0	4
Organic fine chemistry	3.77	2	4.03	0	15
Macromolecular chemistry, Polymers	16.00	14	8.17	4	32
Pharmaceuticals, Cosmetics	3.47	4	2.68	0	8
Biotechnology	0.00	0	0.00	0	0
Agriculture, Food chemistry	0.07	0	0.26	0	1
Chemical and Petrol industry	11.16	10	5.49	4	22
Chemical engineering	1.35	1	0.87	0	3
Surface technology, Coating	3.48	3	2.82	0	9
Materials, Metallurgy	2.41	2	2.12	0	6
Materials processing, Textiles, Paper	3.92	3	2.73	1	9
Handling, Printing	20.26	16	13.55	4	50
Agricultural and Food processing,	0.35	0	0.71	0	2
Environmental technology	3.23	0	4.73	0	15
Machine tools	1.91	1	1.57	0	5
Engines, Pumps and Turbines	21.72	15	21.10	3	69
Thermal processes and apparatus	0.37	0	0.62	0	2
Mechanical elements	2.33	2	2.14	0	7
Transport	16.54	14	12.00	2	50
Space technology, Weapons	0.00	0	0.00	0	0
Consumer goods	0.72	0	1.05	0	4
Civil engineering, Building, Mining	0.00	0	0.00	0	0

4 An Empirical analysis of patenting behaviour

In this section we derive an empirical model with which to test the theory advanced in Section 2. Then, we describe the sample and provide results.

4.1 An empirical model of patenting behaviour

Building on the results of section 2 we estimate a model predicting the level of patent applications filed by a firm in a given year at the EPO. It is well known that firms' patent applications are highly persistent, reflecting long term investments in R&D capacity. To deal with this we include a lagged dependent variable in our model and estimate the following relationship:

$$\begin{aligned} A_{i,t} = \beta_0 + \beta_A A_{i,t-1} + \beta_O O_{i,t} + \beta_C C_{i,t} + \boldsymbol{\beta}_X{}' \boldsymbol{X}_{i,t} \\ + \beta_{AC} A_{i,t-1} C_{i,t} + \beta_{OC} O_{i,t} C_{i,t} + \beta_{OCL} O_{i,t} C_{i,t} L_{i,t} + \beta_{OL} O_{i,t} L_{i,t} \\ + \chi_i + \zeta_{it} \;, \qquad \text{where:} \end{aligned} \tag{18}$$

$A_{i,t}$ – Patent Applications $\quad$ $O_{i,t}$ – Technological Opportunity: Non Patent References

$C_{i,t}$ – Complexity: Triples $\quad$ $\boldsymbol{X}_{i,t}$ – Control variables: Fragmentation, Area count, Size(L)

This specification allows us to control for effects of technological opportunity β_O and complexity β_C in discrete technologies as well as in complex technologies through the interaction terms. We also include interaction terms that allow us to distinguish the patenting behaviour of large and small firms in complex and discrete technologies. We do this as our theoretical model indicates that firms' patenting behaviour will depend on the share of patents they expect to receive on a given technological opportunity.

The empirical model allows us to test the following hypotheses that reflect the predictions derived from the theoretical model:

H1 : Increased technological opportunity lowers the level of patent applications in complex technologies (Proposition 2);

H2 : Increased complexity of a technology raises the level of patent applications in a complex technologies (Proposition 3);

H3 : Increased technological opportunity raises the level of patent applications in discrete technologies (Proposition 4).

Hypothesis 1 implies that $\beta_O + \beta_{OC} \times C_{i,t} < 0$, hypothesis 2 implies that $\beta_C + \beta_{OC} \times O_{i,t} + \beta_{AC} \times A_{i,t} > 0$ and hypothesis 3 implies that $\beta_O > 0$.

4.2 Data and descriptive statistics

Our dataset consists of $173,448$ observations of firms patenting in specific technology areas in a given year. Our data covers the period between 1978 when the EPO began operating and 2003. We excluded patentees from the sample in two steps: first, we excluded all those patentees with fewer than 10 patent applications in a given technology area over the entire period, second we excluded those patentees who had fewer than three years of positive patent applications in a technology area in the fifteen years after 1987. These steps were necessary to exclude firms that cannot be considered to have a long term patenting strategy. Only patentees with a longer patenting horizon will be affected by changes in technological opportunity, or the degree of blocking over time.

Table 2 shows that most firms remaining in the sample have a low annual level of patent applications per area (5.43) but are active in 8 or 9 different technology areas. The large dummy splits firms almost exactly into the largest and smallest firms in the sample. The median technology area contained 5 Triples in a given year. The level of non patent references in the average technology area is 1.151. The table also contains information about sample statistics for the year 1992, after which patent applications increased markedly as Figure 3 shows.

A comparison of sample means and means for 1992 shows that firms patent in more areas, face more Triples and generate fewer non patent references after 1992 than before. This confirms what we have shown graphically in the previous section.

Table 2: Descriptive statistics for the sample (1987-2002)

Variable	**Aggregation level**	**Mean**	**Median**	**Standard deviation**	**Minimum**	**Maximum**
Patent applications	Firm	5.431	1.000	18.594	0.000	752.000
$\log$ Patent applications	Firm	1.051	0.693	1.052	0.000	6.624
Areas	Firm	8.751	7.000	6.027	0.000	30.000
Large dummy	Firm	0.504	1.000	-	0.000	1.000
Non Patent References	Area	1.151	0.894	0.827	0.174	4.532
Triples	Area	18.480	5.000	30.085	0.000	166.000
Fragmentation	Area	0.001	0.000	0.006	0.000	0.355

Observations = 173,448

Sample statistics for 1992

Patent applications	Firm	4.235	1.000	14.024	0.000	387.000
$\log$ Patent applications	Firm	0.923	0.693	0.990	0.000	5.961
Areas	Firm	7.746	6.000	5.563	0.000	27.000
Large dummy	Firm	0.438	0.000	-	0.000	1.000
Non Patent References	Area	1.205	0.970	0.747	0.290	3.554
Triples	Area	15.761	3.000	25.348	0.000	104.000
Fragmentation	Area	0.001	0.000	0.006	0.000	0.168

Observations = 11,325

Table 3 provides additional information about the structure of the panel which we employ. It shows that in total there are 2074 different firms in the dataset. Each firm applied for an average of 628 patents across all areas and years included in the sample. The lower half of the table shows that our sample contains 55.8% of all patents applied for at the EPO and 24.8% of patentees.

We treat firms operating in several technology areas as separate observations in each area. We do this to control for area specific patenting behaviour of individual firms. Where we use panel data, the panel is unbalanced due to entry and exit of firms into technology areas.

Table 3: Panel descriptives for the sample

Firm level (n=2074)	**Mean**	**Median**	**SD**
Total patents	628.27	205	1944.94
Total patents (annual)	37.02	12	111.65
Technological areas (annual)	5.54	4	4.56
Area-Year level (n=650)	**Mean**	**Median**	**SD**
Total patents in area	2594.23	2310	1778.87
Total patents in sample	1449.35	1012	1695.86
Total firms in area	1077.62	893	668.14
Total firms in sample	266.84	263	253.71
Triples	14.67	2	27.69
Non Patent References	0.98	0.75	0.75
Fragmentation	0.001	0	0.009

4.3 Results

In this section we set out empirical results. First, we provide results from fixed effects and ordinary least squares estimation. These results are known to be biased as we include a lagged dependent variable in our model. However, they provide lower and upper bounds on the values of the lagged dependent variable for GMM Bond (2002). We also provide results from dynamic panel estimates using the system GMM estimator with orthogonal deviations as proposed by Arellano and Bover (1995).

We do not reject the predictions of our theoretical model. In particular, we find that greater technological opportunity reduces firms' patenting in complex technologies. This effect is diminished for very large firms but remains present. We also find that growing complexity of blocking has the effect of increasing firms' patenting activity. Additionally, we show that there is weak evidence to suggest that fragmentation of patent ownership affects firms incentives to patent in complex technologies once we control for effects of complexity and technological opportunity.

We estimate OLS and fixed effects models (Table 4) to provide a baseline against which to judge the results of our GMM models (Table 5). Table 4 includes simple models as well as models that test our hypotheses through the interaction of *Triples* with *Non patent references* (NPR). *Triples* measure complexity of blocking while *Non patent references* capture the level of technological opportunity in a technology area at a given time.

Table 4: Coefficients for Simple Models of Patent Applications

	OLS models			Fixed effects models		
Variable	OLS 1	OLS 2	OLS 3	FE 1	FE 2	FE 3
log Patentcount$_{t-1}$	0.599***	0.583***	0.583***	0.172***	0.157***	0.156***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
log Patentcount$_{t-1}$ $\times$ Triples		0.001***	0.001***		0.001***	0.001***
		(0.000)	(0.000)		(0.000)	(0.000)
Non Patent References (NPR)	0.064***	0.076***	0.067***	0.002	0.016	-0.007
	(0.002)	(0.002)	(0.003)	(0.007)	(0.008)	(0.009)
NPR $\times$ Triples		-0.002***	-0.002***		0.000	0.000
		(0.000)	(0.000)		(0.000)	(0.000)
NPR $\times$ Triples $\times$ Large			0.000*			0.000
			(0.000)			(0.000)
NPR $\times$ Large			0.020***			0.038***
			(0.004)			(0.006)
Fragmentation	29.910***	30.332***	30.352***	34.246***	33.811***	33.825***
	(0.269)	(0.320)	(0.320)	(0.346)	(0.392)	(0.392)
Fragmentation $\times$ Triples		-0.028***	-0.028***		0.016	0.016
		(0.007)	(0.007)		(0.009)	(0.009)
Triples	0.000***	0.002***	0.002***	0.001***	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Areas	0.018***	0.018***	0.018***	0.084***	0.084***	0.084***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Large	0.279***	0.282***	0.256***	0.305***	0.306***	0.263***
	(0.004)	(0.004)	(0.006)	(0.005)	(0.005)	(0.009)
Year dummies	YES	YES	YES	YES	YES	YES
Primary area dummies	YES	YES	YES	YES	YES	YES
Constant	0.122***	0.116***	0.128***	0.029	0.031*	0.060***
	(0.011)	(0.011)	(0.012)	(0.015)	(0.016)	(0.016)
R-squared	0.671	0.672	0.672	0.300	0.301	0.301
N	173448	173448	173448	173448	173448	173448

*p<0.05, ** p<0.01, *** p<0.001

Table 5: Coefficients for System GMM Models of Patent Applications

Variable	Allowing correlation with fixed effects				Assuming no correlation with fixed effects	
	SGMM MIN	SGMM END	DGMM END	SGMM END2	SGMM NPR	SGMM F
log Patentcount$_{t-1}$	0.684***	0.678***	0.863***	0.735***	0.715***	0.915***
	(0.072)	(0.068)	(0.091)	(0.058)	(0.047)	(0.039)
log Patentcount$_{t-1} \times$ Triples	-0.017***	-0.015***	-0.012***	-0.011***	-0.007***	-0.004***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Non Patent References (NPR)	1.581***	1.386***	1.198***	0.968***	0.271***	0.171
	(0.221)	(0.182)	(0.164)	(0.113)	(0.019)	(0.119)
NPR $\times$ Triples	-0.038***	-0.034***	-0.028***	-0.024***	-0.008***	-0.003
	(0.005)	(0.004)	(0.004)	(0.002)	(0.001)	(0.003)
NPR $\times$ Triples $\times$ Large	0.006***	0.006***	0.006***	0.005***	0.004***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
NPR $\times$ Large	-0.436***	-0.425***	-0.262***	-0.397***	-0.466***	-0.506***
	(0.055)	(0.052)	(0.033)	(0.042)	(0.034)	(0.032)
Fragmentation	-15.234*	-12.482*	-13.998*	-4.848	-1.448	-2.313
	(6.510)	(6.192)	(6.123)	(3.654)	(1.210)	(1.946)
Fragmentation $\times$ Triples	0.262**	0.247*	0.181*	0.188*	0.102*	0.156*
	(0.100)	(0.097)	(0.091)	(0.083)	(0.044)	(0.071)
Triples	0.063***	0.057***	0.042***	0.040***	0.015***	0.007
	(0.007)	(0.006)	(0.005)	(0.004)	(0.001)	(0.004)
Areas	0.095***	0.096***	0.031*	0.086***	0.085***	0.058***
	(0.010)	(0.010)	(0.014)	(0.008)	(0.007)	(0.006)
Large	0.430***	0.409***	0.257***	0.325***	0.442***	0.412***
	(0.087)	(0.081)	(0.053)	(0.065)	(0.049)	(0.048)
Year dummies	YES	YES	YES	YES	YES	YES
Primary area dummies	YES	YES	YES	YES	YES	YES
Constant	-1.672***	-1.515***		-1.151***	-0.597***	-0.526***
	(0.198)	(0.167)		(0.105)	(0.046)	(0.106)
N	173448	173448	171380	173448	173448	173448
m1	-12.75267	-13.49454	-9.115675	-16.66536	-20.32686	-28.27661
m2	4.690134	5.564835	5.686894	9.293913	12.525	20.07668
m3	1.093296	.7390595	-.4191068	-.4131314	-1.354271	-1.478497
Hansen	2.178791	10.67657	7.067067	70.62775	184.0212	288.5185
Degrees of freedom	1	5	4	9	7	7

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

1. Asymptotic standard errors, asymptotically robust to heteroskedasticity are reported in parentheses
2. m1-m3 are tests for first- to third-order serial correlation in the first differenced residuals.
3. Hansen is a test of overidentifying restrictions. It is distributed as χ^2 under the null of instrument validity, with degrees of freedom reported below.

4. In all cases GMM instrument sets were collapsed and lags were limited.

We also report specifications that allow the effect of *Non patent references* and of *Triples* to vary according to size of the firm's patent portfolio. The OLS models support Proposition 2: greater technological opportunity (Non Patent References) reduces firms' patenting efforts in complex technology areas. In contrast there is no clear result for 3. The fixed effects models provide no clear evidence on either Proposition. Note that in case of the fixed effects models the ratio of $\sigma_\eta^2/\sigma_\nu^2$ is 0.81. It has been shown that the finite sample bias of the system GMM estimator is large if this ratio is large, which is not the case here (Bun and Kiviet (2006),Hayakawa (2006)).

Table 5 presents results of six models estimated with system and difference GMM using orthogonal deviations (Arellano and Bover (1995)).[15] The models presented differ in the number of overidentifying restrictions employed as well as assumptions about the correlation of the explanatory variables with fixed effects. The four models reported in the central part of the table allow for correlation between all explanatory variables apart from *Triples* with fixed effects. In the two specifications on the right side of the table we assume that subsets of the explanatory variables are uncorrelated with fixed effects.

Note that the number of observations in our dataset implies that $T/N \rightarrow 0$. This implies that a systems GMM estimator (Blundell and Bond (1998)) using forward deviations is asymptotically consistent (Alvarez and Arellano (2003); Hayakawa (2006)). We employ this estimator as the persistence of the patenting series is very high in our sample: the coefficient on the lagged dependent variable in an AR1 model with time and primary area dummies is 0.92. Blundell and Bond (1998) note that a difference GMM estimator will be affected by a weak instruments problem in this context. Specification DGMM END reported in Table 5, which is estimated by difference GMM, does not suggest that the weak instruments problem is severe here. However, the coefficient on the lagged dependent variable is somewhat above that reported for the comparable systems estimators. It is also significantly above the coefficients from the OLS regressions reported in Table 4. Therefore, we focus our analysis on the results from the system estimators. Note however that the substantive results provided by the difference estimator are the same as those from the systems estimators.

In all models reported in Table 5 the instrument sets were collapsed[16] and instrumenting lags were limited as described below. This was done as the Hansen test and difference in Hansen tests rejected the overall instrument sets as well as individual instruments where larger instrument sets were employed. Specification SGMM END2 illustrates how sensitive the Hansen test is to the size of the instrument set here. This specification is identical to SGMM END3, we just allow for an extra lag on the instrument sets for the endogenous variables in this specification. The specification is rejected by the Hansen test at conventional significance levels.

All models reported in Table 5 contain the following explanatory variables: *Non patent references*, *Triples*, *Fragmentation*, *Area count*, *Large dummy* and the lagged dependent variable as well as interactions of some of these variables. We consider *Large* and *Area count* to be endogenous as they reflect decisions about how widely and where to engage in research which may be contemporaneous with decisions determining the level of patent applications. We consider the remaining variables to be predetermined since they depend in large part on the aggregated decisions of rival firms. Finally note that we include only year and primary area dummies as well as *Triples* in the levels equation as

[15] All models were estimated with xtabond2 in Stata 9.2 . This package is described in (Roodman (2006)).

[16] Collapsing instrument sets reduces the number of moment conditions used for GMM (Roodman (2006)).

it is likely that the fixed effects are correlated with differences in the remaining explanatory variables. *Triples* is the only variable that reflects purely technology area specific characteristics which may be assumed to be orthogonal to firm specific effects.

We estimate two models in which we treat Fragmentation (GMM F) and Non patent references (GMM NPR) as uncorrelated with fixed effects. Results from the Hansen tests for both specifications reported in Table 5 show that these models are clearly rejected.

Our preferred models are reported as SGMM MIN and SGMM END in Table 5. In SGMM MIN we restrict the number of instruments such that the model is just overidentified. Hayakawa (2006) argues that such a minimum instruments specification is unbiased in settings where T is fixed and N$\rightarrow \infty$. The specification SGMM END includes one additional lag for the endogenous variables. Results from these two specifications are statistically indistinguishable.

Focusing on these two specifications we find that all our theoretical predictions are borne out by the data. First, we find that in discrete technologies additional technological opportunity raises firms' patenting rates. The coefficient for *Non patent references* is positive and highly significant. Even in case of large firms the overall effect remains positive. This shows that Hypothesis III cannot be rejected. Second, the coefficient on the interaction of *Non patent references* and *Triples* is negative. The overall effect of additional Non patent references on patenting becomes negative if there are more than 42 Triples in a technology area. As Table 1 shows this is the case in at least one year for eight of the technology areas in our sample. For Audiovisual technology and Optics it is always the case! In case of larger firms the predicted effects of complexity already arise when the number of Triples is above 4. This is always the case for 9 technology areas in our sample! These findings show that Hypothesis I cannot be rejected in our sample. Finally, the coefficient on *Triples* is positive and greater than that on the sum of interactions of *Triples* with *Non patent references* and . This shows that greater blocking complexity and therefore greater complexity of a technology area increase firms' levels of patenting. Therefore, we cannot reject Hypothesis II.

Additionally, our results indicate that the persistsence of patenting decreases as technology areas become more complex. This suggests that patentees are more responsive to their competitors' patenting behaviour in complex technology areas than in discrete technology areas. Finally, our results show that fragmentation of patent portfolios increases firms' patenting efforts in technology areas in which there are more than 50 Triples, i.e. if blocking complexity is high. This effect confirms the findings of Ziedonis (2004). We also find that fragmentation of patent ownership in discrete technology areas has negative effects on firms' patenting efforts.

5 Conclusion

Patent applications have been increasing steeply at the USPTO and the EPO since 1984 and 1992 respectively. In both cases these increases have raised questions about the operations of the affected patent offices as well as effects of these trends on economic activity more generally (F.T.C. (2003); **?** and von Graevenitz et al. (2007)). The increases in patenting are concentrated in complex technologies (Hall (2005) and von Graevenitz et al. (2007)) which are prone to the formation of patent thickets. Resulting increases in transactions costs affect technologies that are central to the large productivity increases recorded in the recent past Jorgenson and Wessner (2007).

There is a great deal of evidence that patenting has increased in response to evolution of the legal environment, specifically in the United States, changes in the management of R&D and patenting, increasing complexity of technology and more strategic behaviour of patent applicants (Kortum and Lerner (1998); Hall and Ziedonis (2001); Ziedonis (2004)). The contribution of technological fecundity to current patenting trends is less well understood.

We provide a model of patenting that encompasses discrete and complex technologies. The model captures effects of technological opportunity and technological complexity on firms' patenting incentives in contexts in which firms' interact strategically through the patent system. We show that greater technological opportunity will raise patenting in discrete technologies but will lower it in complex technologies. We also show that greater complexity of technologies will raise firms' patenting levels.

Using data on patenting in Europe we show that our model predicts firms' patenting behaviour well. All predictions of the theoretical model are confirmed. We capture the dynamic aspects of patenting behaviour by including a lagged dependent variable in our empirical specifications. In order to control for the endogeneity this creates we estimate our models using GMM estimators (Arellano and Bover (1995); Blundell and Bond (1998); Alvarez and Arellano (2003)).

In order to estimate our model we derive a new measure of blocking complexity which is derived from patent data. This measure captures the extent to which patent thickets have formed in specific technology areas. The measure also shows us how complex different technologies are. With the help of this measure we show that patent thickets exist in nine out of thirty technology areas at the EPO.

Our data indicate that the problem of patent thickets at the EPO is getting worse in recent years. In future work we intend to investigate whether this has measurable effects on the productivity of firms' R&D investments.

References

ALVAREZ, J. AND M. ARELLANO (2003): "The Time Series and Cross-Section Asymptotics of Dynamic Panel Data Estimators," *Econometrica*, 71, 1121–1159.

AMIR, R. (2005): "Supermodularity and Complementarity in Economics: An Elementary Survey," *Southern Economic Journal*, 71, 636–660.

ARELLANO, M. (2003): *Panel Data Econometrics*, Advanced Texts in Econometrics, Oxford, GB: Oxford University Press.

ARELLANO, M. AND O. BOVER (1995): "Another Look at the Instrumental-Variable Estimation of Error-Components Models," *Journal of Econometrics*, 68, 29–51.

ARUNDEL, A. (2001): "The Relative Effectiveness of Patents and Secrecy for Appropriation," *Research Policy*, 30, 611–624.

——— (2003): "Patents in the Knowledge-Based Economy, Report of the KNOW Survey," Tech. rep., MERIT, University of Maastricht.

ARUNDEL, A. AND I. KABLA (1998): "What Percentage of Innovations are Patented? Empirical Estimates form European Firms," *Research Policy*, 27, 127–141.

BEATH, J., Y. KATSOULACOS, AND D. ULPH (1989): “Strategic R&D Policy,” *The Economic Journal*, 99, 74–83.

BESSEN, J. (2004): “Patent Thickets: Strategic Patenting of Complex Technologies9,” Tech. rep., Research on Innovation.

BESSEN, J. AND M. J. MEURER (2008): *Patent Failure*, Princeton University Press.

BLUNDELL, R. AND S. BOND (1998): “Initial conditions and moment restrictions in dynamic panel data models,” *Journal of Econometrics*, 115–143.

BOND, S. (2002): “Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice,” CEMMAP Working Paper CWP09/02, IFS.

BUN, M. AND J. KIVIET (2006): “The Effects of Dynamic Feedbacks on LS and MM Estimator Accuracy in Panel Data Models,” *Journal of Econometrics*, forthcoming.

CARTER, M. (2001): *Foundations of Mathematical Economics*, Cambridge, Massachusetts and London, England: MIT Press.

CLARK, D. J. AND K. A. KONRAD (2005): “Fragmented Property Rights, R&D and Market Structure,” Working Paper 4, University of Tromso, Department of Economics and Management.

COHEN, W. M., R. R. NELSON, AND J. P. WALSH (2000): “Protecting their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not),” Working Paper 7552, NBER.

GALLINI, N. T. AND S. SCOTCHMER (2002): “Intellectual Property: When Is It the Best Incentive System?” in *Innovation Policy and the Economy*, ed. by J. Lerner, S. Stern, and A. B. Jaffe, Cambridge, Mass.: MIT Press, vol. 2.

GRINDLEY, P. C. AND D. J. TEECE (1997): “Managing Intellectual Capital: Licensing and Cross-Licensing in Semiconductors and Electronics,” *California Management Review*, 39, 8–41.

HALL, B. H. (2005): “Exploring the Patent Explosion,” *Journal of Technology Transfer*, 30, 35–48.

HALL, B. H. AND D. HARHOFF (2004): “Post Grant Review Systems at the U.S. Patent Office - Design Parameters and Expected Impact,” *Berkeley Technology Law Journal*, 19, 981–1015.

HALL, B. H. AND R. ZIEDONIS (2001): “The Patent Paradox Revisited: An Empirical Study of Patenting in the U.S. Semiconductor Industry, 1979-1995,” *Rand Journal of Economics*, 32, 101–128.

HARHOFF, D., K. HOISL, AND C. WEBB (2006): “European Patent Citations - How to count and how to interpret them,” .

HAYAKAWA, K. (2006): “The Asymptotic Properties of the System GMM Estimator in Dynamic Panel Data Models When Both N and T are Large,” HiStat Discussion Paper Series 129, Hitotsubashi University.

JAFFE, A. B. (2000): “The U.S. Patent System in Transition: Policy Innovation and the Innovation Process,” *Research Policy*, 29, 531–557.

JORGENSON, D. W. AND C. W. WESSNER, eds. (2007): *Enhancing Productivity Growth in the Information Age: Measuring and Sustaining the New Economy*, National Research Council.

KORTUM, S. AND J. LERNER (1998): “Stronger Protection or Technological Revolution: What is Behind the Recent Surge in Patenting?” in *Carnegie-Rochester Conference Series on Public Policy*, vol. 48, 247–304.

——— (1999): “What is behind the Recent Surge in Patenting?” *Research Policy*, 28, 1–22.

KUSONAKI, K., I. NONAKA, AND A. NAGATA (1998): “Organizational capabilities in product development of Japanese firms,” *Organization Science*, 9, 699–718.

LEE, T. AND L. L. WILDE (1980): “Market Structure and Innovation: A Reformulation,” *The Quarterly journal of Economics*, 94, 429436.

LEVIN, R. C., A. K. KLEVORICK, R. R. NELSON, AND S. G. WINTER (1987): “Appropriating the Returns from Industrial Research and Development,” *Brookings Papers on Economic Activity*, 3.

MEYER, M. (2000): “Patent Citations in a Novel Field of Technology - What can they tell about Interactions Between Emerging Communities of Science and Technology?” *Scientometrics*, 48, 151–178.

MILGROM, P. AND J. ROBERTS (1990): “Rationalizability, learning and equilibrium in games with strategic complementarities.” *Econometrica*, 58, 1255–1278.

NARIN, F., K. HAMILTON, AND D. OLIVATRO (1997): “The Increasinf Linkage Between U.S. Technology and Public Science,” *REsearch Policy*, 26, 317–330.

NARIN, F. AND E. NOMA (1985): “Is Technology becoming Science,” *Scientometrics*, 7, 369–381.

OECD (1994): “Using Patent Data as Science and Technological Indicators,” Tech. rep., Patent Manual.

F.T.C. (2003): “To Promote Innovation - The Proper Balance of Competition and Patent Law and Policy,” Report, GPO.

NATIONAL RESEARCH COUNCIL (2004): *Patent System for the 21st Century, Report of the Board on Science, Technology, and Economic Policy*, Washington, D.C.: National Academies Press.

REINGANUM, J. F. (1989): “The Timing of Innovation: Research, Development and Diffusion,” in *Handbook of Industrial Organization*, ed. by R. Schmalensee and R. D. Willig, Amsterdam, Netherlands: Elsevier Science, vol. 1, 850–908.

ROODMAN, D. (2006): “How to Do xtabond2: An Introduction to ”Difference” and ”System” GMM in Stata,” Working Paper 103, Center for Global Development, Washington.

RUBINFELD, D. L. AND R. MANESS (2005): “The Strategic Use of Patents: Implications for Antitrust,” in *Antitrust, Patents and Copyright*, ed. by F. Lvque and H. Shelanski, Edward Elgar.

SCHANKERMAN, M. AND M. NOEL (2006): “Strategic Patenting and Software Innovation,” Discusion Paper 5701, CEPR.

SCOTCHMER, S. (2005): *Innovation and Incentives*, Cambridge, Mass.: MIT Press.

SHAPIRO, C. (2001): “Navigating the Patent Thicket: Cross Licenses, Patent Pools, and Standard-Setting,” in *Innovation Policy and the Economy*, ed. by A. B. Jaffe, J. Lerner, and S. Stern, Cambridge, Mass.: MIT Press, vol. 1.

——— (2003): “Technology Cross-Licensing Practices: F.T.C. vs. Intel (1999),” in *The Antitrust Revolution*, ed. by J. E. Kwoka and L. J. White, Oxford: Oxford University Press, 350–369.

SIEBERT, R. AND VON GRAEVENITZ (2006): “Jostling for advantage - what does it take to make partners compete ?” .

SIEBERT, R. AND G. VON GRAEVENITZ (2008): “Does Licensing Resolve Hold Up in the Patent Thicket?” Discussion Paper 01, Munich School of Management.

SOMAYA, D., I. O. WILLIAMSON, AND X. ZHANG (2007): “Combining Patent Law Expertise with RD for Patenting Performance,” *Organization Science*.

VIVES, X. (1990): “Nash equilibrium with strategic complementarities,” *Journal of Mathematical Economics*, 19, 305–321.

——— (1999): *Oligopoly Pricing, Old Ideas and New Tools*, Cambridge, Massachusetts and London, England: MIT Press.

VON GRAEVENITZ, G., S. WAGNER, K. HOISL, B. HALL, D. HARHOFF, P. GIURI, AND A. GAMBARDELLA (2007): “The Strategic Use of Patents and its Implications for Enterprise and Competition Policies,” Tech. rep., European Commission.

ZIEDONIS, R. H. (2004): “Don’t Fence Me In: Fragmented Markets for Technology and the Patent Acquisition Strategies of Firms,” *Management Science*, 50, 804820.

Appendix

A Complex and discrete technologies

Table 6: Classification of technology areas according to OST-INPI/FhG-ISI

Area Code	Description	Classification
1	Electrical machinery, electrical energy	Complex
2	Audiovisual technology	Complex
3	Telecommunications	Complex
4	Information technology	Complex
5	Semiconductors	Complex
6	Optics	Complex
7	Analysis, measurement, control technology	Complex
8	Medical technology	Complex
9	Nuclear engineering	Complex
10	Organic fine chemistry	Discrete
11	Macromolecular chemistry, polymers	Discrete
12	Pharmaceuticals, cosmetics	Discrete
13	Biotechnology	Discrete
14	Agriculture, food chemistry	Discrete
15	Chemical and petrol industry, basic materials chemistry	Discrete
16	Chemical engineering	Discrete
17	Surface technology, coating	Discrete
18	Materials, metallurgy	Discrete
19	Materials processing, textiles paper	Discrete
20	Handling, printing	Discrete
21	Agricultural and food processing, machinery and apparatus	Discrete
22	Environmental technology	Complex
23	Machine tools	Complex
24	Engines, pumps and turbines	Complex
25	Thermal processes and apparatus	Complex
26	Mechanical elements	Complex
27	Transport	Complex
28	Space technology, weapons	Complex
29	Consumer goods and equipments	Complex
30	Civil engineering, building, mining	Complex

Description of the 30 technology areas contained in the OST-INPI/FhG-ISI technology nomenclature. We classified the 30 technology areas as complex or discrete attempting to replicate the classification of Cohen et al. (2000).

Type	Description
X	Particularly relevant documents when taken alone (a claimed invention cannot be considered novel or cannot be considered to involve an inventive step)
Y	Particularly relevant if combined with another document of the same category
A	Documents defining the general state of the art
O	Documents referring to non-written disclosure
P	Intermediate documents (documents published between the date of filing and the priority date)
T	Documents relating to theory or principle underlying the invention (documents which were published after the filing date and are not in conflict with the application, but were cited for a better understanding of the invention)
E	Potentially conflicting patent documents, published on or after the filing date of the underlying invention
D	Document already cited in the application
L	Document cited for other reasons (e.g., a document which may throw doubt on a priority claim)

Table 7: Overview over different types of references included in the search report by the EPO. Source: EPO Guidelines for Examination in the European Patent Office, 2003.