The propensity to patent:
An empirical analysis at the innovation level[*]

Iiro Mäkinen
ETLA – The Research Institute of the Finnish Economy

A paper submitted to the EPIP-2007 Conference
Lund, Sweden, September 20-21, 2007

Abstract

This study seeks to shed new light on the complex relationship between innovations and patents. The objective of the study is to contribute to our understanding of which innovations are patented—and which are not—by analyzing the patenting decision for circa 800 Finnish product innovations. The data is drawn from the Sfinno database compiled at VTT Technical Research Centre of Finland. The econometric analysis indicates that various characteristics of the innovation, the market, and the innovating firm have a significant effect on the propensity to patent. First, there appears to be a U-shaped relationship between firm size and the propensity to patent, which can be attributed to a relatively large extent to economies of scale in the patenting activity as well as to the relatively important role of patenting in start-up ventures. Second, the estimation results suggest that larger—that is, more novel and significant—innovations are patented more frequently than smaller ones. Third, technologically very complex innovations appear to be patented less often than others, while the fragmentation of intellectual property rights to cumulatively developing technology seems to entail high propensities to patent.

Keywords: patents, innovations, patenting, propensity.

---

[*] This paper is based on the author's Master's thesis (Mäkinen, 2007) submitted to Helsinki School of Economics (HSE). The thesis work was carried out within the CIPCI project at VTT Technical Research Centre of Finland. Financial support from VTT and Tekes is gratefully acknowledged. The paper has benefited greatly from insightful comments by the Master's thesis supervisor Heli Koski from ETLA, Professor Pekka Ilmakunnas from HSE, and the CIPCI project manager Jani Saarinen and others from VTT Innovation Studies. All responsibility for any errors and omissions is solely mine.
Correspondence: Iiro Mäkinen, ETLA – The Research Institute of the Finnish Economy, Lönnrotinkatu 4 B, FIN-00120 Helsinki, Finland. Email: iiro.makinen@etla.fi.

1. Introduction

To patent or not to patent: that is the question innovators face when they succeed in developing novel products or processes. The innovators need to contemplate whether it is better to seek patent protection or strive to appropriate returns to innovation through other means such as secrecy, first-mover advantages, and complementary capabilities. Various scholars have noted that the propensity to patent differs across industries, firms, and kinds of innovations (e.g. Basberg, 1987; Griliches, 1990; Patel and Pavitt, 1995; Archibugi and Pianta, 1996; Kleinknecht et al., 2002). However, precious little is known about the origins of such differences, especially at the level of innovations, and several issues remain ambiguous in both theoretical and empirical literature.

In the theoretical economic literature on patents, the patenting decision is modeled as a profit-maximizing choice between patenting and non-patenting strategies[1] (e.g. Horstmann et al., 1985; Scotchmer and Green, 1990; Waterson, 1990; Gallini, 1992; Takalo, 1998; Denicolò and Franzoni, 2003; Anton and Yao, 2004; Kultti et al., 2007). This literature is primarily concerned with the optimal design and welfare effects of the patent system on a very general level. Hence most theoretical models abstract from the heterogeneity of industries, firms, and innovations, and provide relatively little insight into the determinants of the propensity to patent. And when relevant predictions emerge from the theoretical work, they can be very sensitive to the assumptions of the specific models. The Anton and Yao (2004) model, for instance, implies that small innovations are patented while large innovations are kept secret, whereas the Horstmann et al. (1985) and the Denicolò and Franzoni (2003) models arrive at the opposite conclusion[2].

The empirical studies on the propensity to patent have hitherto been generally confined to the use of industry and firm-level data, and thus we have very little idea of how the propensity to patent varies across different innovations. Moreover, due to different and sometimes problematic definitions of the propensity to patent in these studies, the results are not readily comparable. And when comparisons are attempted, contradictory conclusions seem to emerge. The results of Schmookler (1966), Taylor and Silberston (1973), and Bound et al. (1984), for instance, suggest that the propensity to patent decreases with the scale of operations, while Mansfield (1986), Arundel and Kabla (1998), Duguet and Kabla (1998), and Arora et al. (2003) find support for the opposite conclusion. Hence further empirical research is required to broaden and deepen our understanding of the determinants of the propensity to patent.

The variations in the propensity to patent are not a trivial matter, but they do have important implications for researchers and policy makers with an interest in innovation and technological change. The patent system is an important policy instrument that can be used to affect the allocation of resources for innovative activities and the diffusion of innovations. Variations in the propensity to patent can be indicative of differences in the extent to which the patent system is utilized by different firms to appropriate returns to different innovations. Furthermore, a thorough understanding of the variations in the propensity to patent should be of great value to researchers, policy makers, and others who depend on patent data in drawing conclusions about innovation and technological change.

The fact that not all innovations are patented is often pointed out as a major limitation to the use of patent statistics as an indicator of innovation (e.g. Griliches, 1990; Archibugi and Pianta, 1996; Kleinknecht et al., 2002). As Hall et al. (2001:4) point out:

> “Unfortunately, we have very little idea of the extent to which patents are representative of the wider universe of inventions, since there is no systematic data about inventions that are not patented. This is an important, wide-open area for future research.”

Whether small innovations are patented while large ones are kept secret, or vice versa, should have major implications for the utilization of patent data in economic research. Moreover, understanding the relationship between firm size and the propensity to patent is essential in interpreting empirical studies on the Schumpeterian hypotheses[3] that use patents as a measure of innovation.

The objective of this study is to contribute to our understanding of which innovations are patented—and which are not—by analyzing the patenting decision for circa 800 Finnish product innovations contained in a unique innovation database compiled at VTT Technical Research Centre of Finland. With the help of econometric methods, this study aims to shed new light on the following question: *How is the propensity to patent an innovation affected by the characteristics of the innovation, the market, and the innovating firm?*

This paper is structured as follows. Section 2 discusses issues related to the definition and measurement of the propensity to patent. Section 3 lays out the background for the innovation-level analysis by introducing the data and outlining the hypotheses to be addressed in the empirical study. Section 4 presents the econometric modeling and the estimation results. Section 5 concludes.

2. The propensity to patent: definition and measurement

The relationship between ideas, innovations, and patents is not as clear and simple as it appears in the theoretical literature. Ideally, a firm encounters an idea—or an investment opportunity—and decides whether it is worthwhile investing in developing the idea into an innovation. And if an innovation is successfully developed, the firm then decides whether the innovation should be patented. (Cf. Gallini, 1992; Takalo, 1998; Kultti et al., 2007.) In such a stylized context the definition of the propensity to patent as the fraction of innovations that are patented is straightforward and unambiguously defines the relationship between innovations and patents. In reality, however, it is possible that inventions that are not successfully implemented into practice—and thus do not qualify as innovations[4]—are nevertheless patented. On the other hand, not all inventions are patentable even if they are successfully introduced to the market. It can also happen that the innovator decides to patent but the patent examiner deems the innovation unpatentable and denies the application. Figure 1 illustrates the relationship between ideas, inventions, innovations, and patents. Furthermore, a single innovation can sometimes be protected by a myriad of patents, while one patent can protect a set of innovations. This further complicates the relationship between innovations and patents by making a clear-cut one-to-one mapping between innovations and patents impossible. The complexity of the relationship between innovations and patents, together with problems related to the definition and measurement of innovation, give rise to a number of different definitions of the propensity to patent in the empirical literature.

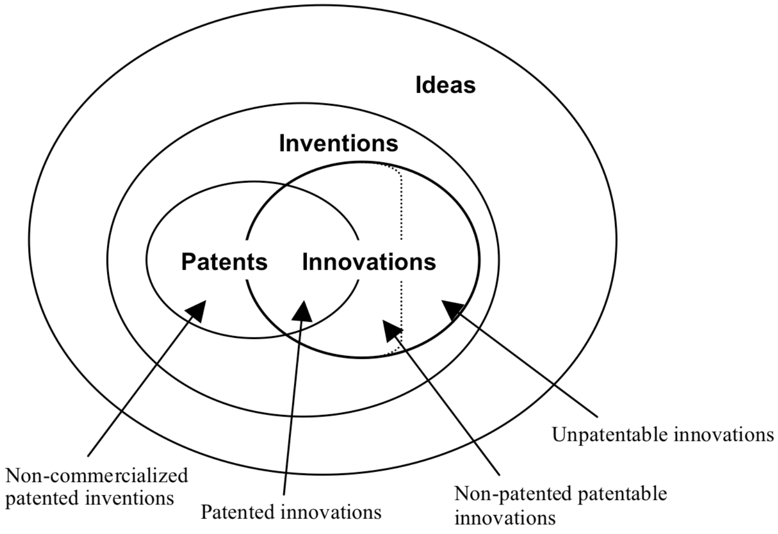

Figure 1. Ideas, inventions, innovations, and patents[5].

Patents, R&D, and the patent production function. Scherer (1965) uses the number of patents received per thousand R&D employees to measure the differences in the propensity to patent, although he acknowledges this to be a crude measure of the patented proportion of the innovation output. Taylor and Silberston (1973) and Scherer (1983) take a relatively similar approach and define the propensity to patent in terms of patents obtained per unit of R&D expenditure. The results on the relationship between patents and R&D are very complex to interpret, however, because they can be affected either by the productivity of R&D or the propensity to patent the results of that R&D[6]. In fact, much of the research on the patents-R&D relationship is primarily concerned with the productivity of R&D, while variations in the propensity to patent are only discussed because they can compromise the interpretability of the results obtained. For instance, it is a matter of speculation whether the negative relationship between the ratio of patents to

R&D and the scale of R&D activities or firm size—observed in a number of studies—arises as a result of declining R&D productivity, decreasing propensity to patent, or something else (e.g. Scherer, 1965, 1983; Schmookler, 1966; Taylor and Silberston, 1973; Bound et al., 1984; Griliches, 1990). In order to distinguish the propensity-to-patent effect from the productivity effects, Brouwer and Kleinknecht (1999) seek to control for the innovation output rather than the innovation inputs by including the sales of innovative products as a control variable in their patent production function.

Survey evidence on the propensity to patent. Instead of seeking to make inferences about the propensity to patent by estimating the patent production function, several innovation surveys have directly asked the firms about the fraction of innovations they generally patent (e.g. Mansfield, 1986; Arundel and Kabla, 1998; Duguet and Kabla, 1998; Cohen et al., 2000; Arora et al., 2003). The survey approach allows for the construction of a direct measure of the propensity to patent that is closely in line with the theoretical definition of the propensity to patent as the fraction of innovations that are patented. Mansfield (1986) defines the propensity to patent as the percentage of patentable inventions that are patented, while the more recent surveys define it as the percentage of innovations for which a patent application is filed[7]. Since in reality a number of patent applications can be filed for a single innovation, the propensity to patent should accordingly be understood as the fraction of innovations for which at least one patent application is filed. This is the definition adopted for the present study.

3. Towards an innovation-level analysis of the propensity to patent

As argued in the introduction, innovation-level data is needed to advance our understanding of the determinants of the propensity to patent. De Melto et al. (1980), Saarinen (2005), and Van der Panne and Kleinknecht (2005) are among the very few studies that have provided innovation-level information on the propensity to patent. These studies, however, do not take the analysis of the propensity to patent very far. De Melto et al. (1980) and Saarinen (2005) address variations in the propensity to patent in the context of Canadian and Finnish innovations, respectively, by cross-tabulating the percentage of innovations patented against other variables of interest. Van der Panne and Kleinknecht (2005) seek to take the analysis a step further by analyzing a sample of Dutch innovations. Their logit analysis, however, is confined by a limited number of observations (N = 216) and explanatory variables (5). The Sfinno database compiled at VTT Technical Research Centre of Finland allows for a detailed innovation-level analysis of the propensity to patent, which simultaneously considers a number of relevant factors hypothesized to affect the patenting decision.

3.1 Sfinno methodology and data

The Sfinno methodology builds upon the object-based method of collecting data on innovative activities directly at the level of individual innovations (cf. Kleinknecht and Bain, 1993). It combines the literature-based method with the expert opinion method in order to produce a comprehensive dataset with a good coverage across different industries and firm size groups (Palmberg et al., 2000). A systematic review of 18 carefully selected trade and technical journals from the period 1985–1998 has been complemented with a review of annual reports of large firms from the same period and with expert opinion-based identification of innovations (Palmberg et al., 1999, 2000; Saarinen, 2005). The review of journals resulted in the identification of some 1100 innovations and the annual reports and expert-opinion yielded about 500 additional innovations giving rise to a dataset of approximately 1600 innovations. In line with the Schumpeterian definitions (Schumpeter 1912) and drawing loosely upon the Oslo Manual (OECD, 1997), the Sfinno approach defines an innovation as an invention that has been commercialized on the market by a business firm or an equivalent, and the inclusion of an innovation in the database requires that the innovation is a technologically new or significantly enhanced product compared to the firm's previous products (Palmberg et al., 1999, 2000). Since the Sfinno-approach relies heavily on public sources in the identification of innovations, it is clearly more conducive to studying product than process innovations. Hence innovations only developed for the firm's internal use are not included in the Sfinno database (Ibid).

In order to collect additional data on the innovations and the development processes, a survey questionnaire was sent to respondents knowledgeable about the innovations in question. Identification of an allegedly relevant respondent was possible for some 1300 innovations and around 800 questionnaires were returned, giving rise to a response rate of over 60 percent (Tanayama, 2002). Moreover, the survey data was complemented with firm-specific data from Statistics Finland and patent data from the National Board of Patents and Registration of Finland. This study is based on a sample of the survey data for which the relevant variables are available. The sample contains 791 innovations from 555 firms. Figure 2 shows the number of firms in the sample with a given number of innovations in the sample. The fact that the data contains several innovations from certain firms suggests that the observations may be subject to within-firm correlation. This issue will be addressed in Section 4.

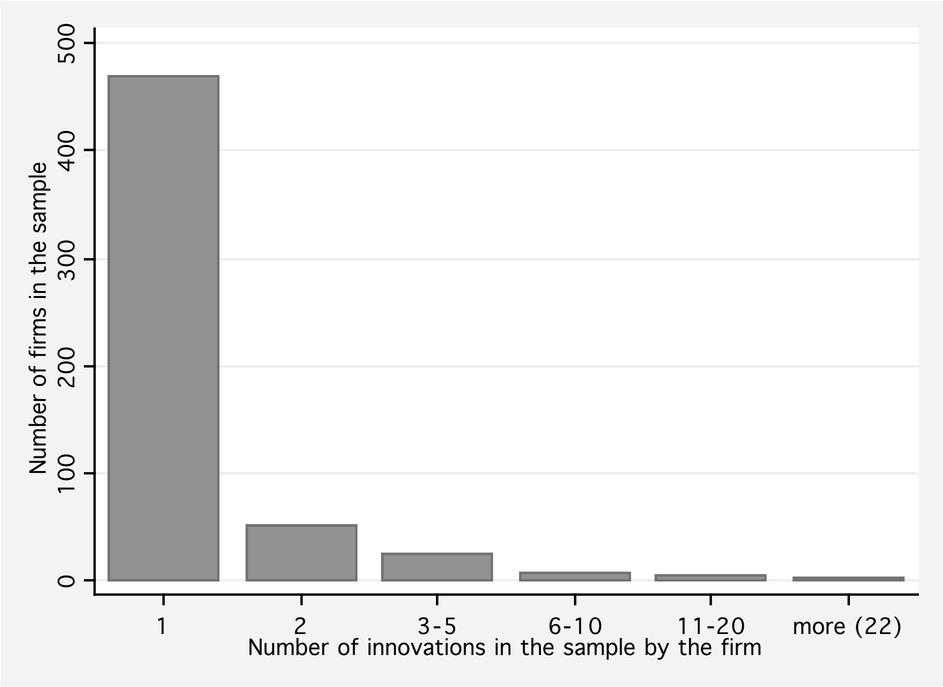

Figure 2. Firms in the sample with a given number of innovations.

An important limitation to innovation-level data collection is that it cannot be based on standard statistical sampling since the population of innovations is unknown (e.g. Palmberg et al., 1999, 2000; Leppälahti, 2000; Palmberg, 2001; Tanayama, 2002; Kleinknecht et al., 2002). Hence, as Tanayama (2002) points out, there is a trade-off between obtaining innovation-level data and collecting data with the desired statistical properties. According to Palmberg (2001:3), data collection in the spirit of the Sfinno approach could instead be described as "a designed census with the aim of identifying all possible products adhering to the specific definition used". Furthermore, Palmberg (2001) argues that "the coverage of the [Sfinno] database in terms of industries and firm size groups is nonetheless relatively representative of innovative activity in Finnish industry" (cf. Leppälahti, 2000; Palmberg et al., 2000). All in all, it can be argued that the Sfinno database is relatively representative of significant Finnish product innovations.

3.2 Hypotheses on the determinants of the patenting decision

Given the innovation-level nature of the data, the dependent variable for the econometric analysis of the propensity to patent takes the form of a binary variable (PATAPP) indicating whether at least one patent application was filed for the innovation. It is of considerable interest as such that patent protection was sought for less than 60 percent of the 791 relatively significant product innovations contained in the sample. This subsection outlines the main hypotheses on the determinants of the propensity to patent emerging from both empirical and theoretical literature, while also introducing the variables designed to capture these determinants. A summary of the variables is provided in Appendix 1.

Firm characteristics. The relationship between firm size and the propensity to patent has been a subject of interest for quite some time, but the evidence remains inconclusive. Even though recent research suggests a positive relationship (Arundel and Kabla, 1998; Duguet and Kabla, 1998; Arora et al., 2003), there are reasons to believe that the relationship might not be as clear-cut as these studies indicate. For one thing, the smallest firms are missing from most of the studies, while the innovation-level studies of Saarinen (2005) and Van der Panne and Kleinknecht (2005) suggest that very small (young) firms can exhibit high propensities to patent. Moreover, firm size might not be independent of the characteristics of innovations. Reinganum (1983) and Henderson (1993), for instance, demonstrate that entrants can have greater incentives to invest in "sufficiently radical innovations". Conversely, Schmookler (1966:35) argues that "one cannot doubt that the largest-scale inventions are usually attempted in large firms". All in all, it seems clear that the differences in the propensities to patent observed in the firm-level investigations might also reflect differences in the characteristics of innovations, not only some inherent firm size-related patenting propensities. Hence it is of great importance to control for the characteristics of innovations when investigating the impact of firm-level factors on the propensity to patent, and vice versa.

A natural explanation for the positive relationship between firm size and the propensity to patent is that economies of scale exist in patenting due to the fixed cost of maintaining a legal department dealing with intellectual property rights (e.g. Scherer, 1965; Lerner, 1995; Arundel and Kabla, 1998; Duguet and Kabla, 1998; Licht and Zoz, 1998; Cohen et al., 2000; Hall and Ziedonis, 2001). Due to their smaller scale of operations, small firms usually cannot spread the fixed costs of patent acquisition and enforcement over as large a volume of inventions as the large firms. There may also be potential for learning curve benefits in the patenting activity. Lerner (1995), for instance, suggests that firms learn to manage internal and external counsel more efficiently when they accumulate experience of litigation. This gives rise to a significant learning curve in the patent litigation process. All in all, it has been argued that small firms cannot utilize the patent system as efficiently as larger firms because obtaining and enforcing patents can be prohibitively costly for many small firms with minimal patent portfolios (e.g. Cohen et al., 2000; Lanjouw and Schankerman, 2004; Parchomovsky and Wagner, 2005). Lanjouw and Schankerman (2004), for instance, find that the litigation risk declines with the size of the patent portfolio. These considerations give rise to the following hypothesis:

Hypothesis 1a: The propensity to patent increases with the scale of patenting.

Despite the problems that a small firm might experience in obtaining and enforcing patents, there are several reasons why small firms might patent more intensively than others. Levin et al. (1987) and Barnett (2003), for instance, argue that small start-up ventures can be more dependent on the patent system than larger firms, because other means of appropriating returns to innovation, such as first-mover advantages and investment in complementary sales and service efforts, may not be viable alternatives. Similarly, Griliches (1990:1676–1677) suggests that for small firms

> "… patents may represent their major hope for ultimate success and hence would lead them to pursue them with more vigor. A well-established major firm does not depend as much on current patenting for its viability or the survival of its market position. Thus, even at an equal underlying inventiveness rates, the propensity to patent may be lower for large firms, at least relative to the successful new entrants in their field."

Small start-ups may often be unable to commercialize their innovations efficiently in embodied form (Cohen and Klepper, 1996), and they thus seek to exploit their innovative technologies through licensing or through a complete transfer of intellectual property. In such situations patents are important for reducing transaction costs and facilitating trade in immaterial property (Arora et al., 2001). Moreover, patents can play an important role as signals of attributes of the firm and the innovations that are deemed positive by outsiders such as venture capitalists and potential collaborators (e.g. Cohen et al., 2000; Kortum and Lerner, 2000; Long, 2002; Hall, 2005). The need for external funding in start-up ventures can also encourage patenting because in order to attract funding the innovator must usually disclose the details of the innovation (Kortum and Lerner, 2000). This can render secrecy a problematic means for appropriation, making formal property rights such as patents an attractive alternative. Hence the following hypothesis is proposed:

Hypothesis 1b: Start-up ventures exhibit high propensities to patent.

The above discussion implies that the relationship between firm size and the propensity to patent may well be non-monotonic. Hence, dummy variables representing four size classes are used to capture the potentially non-linear relationship between firm size and the propensity to patent. The classes of less than 10, 10–99, 100–999, and 1000 or more employees give rise to dummy variables EMP1, EMP2, EMP3, and EMP4, respectively. In order to disentangle the different size-related effects proposed in Hypotheses 1a and 1b, additional variables for the scale of patenting and the start-up status of the innovator are needed.

An innovating firm is defined as an innovative start-up if the idea for the innovation had arisen before or during the year in which the firm was established. The start-up status is coded as a binary variable (STARTUP). Unfortunately, the construction of a measure for the scale of patenting is somewhat problematic because the data does not contain information on the date a patent application was (possibly) filed for the innovation. Hence it is possible that the decision to patent ends up affecting the variable designed to measure the scale of prior patenting, causing simultaneous causality. A measure that should not be very sensitive to such simultaneous causality is the number of patent applications the firm filed (at the National Board of Patents and Registration of Finland) the year before the development of the innovation started (PATENTS). This is based on the assumption that some development work needs to be undertaken before the original idea can be translated into a patentable application. The problem of simultaneous causality will be assessed in Section 4 by testing the exogeneity assumption of PATENTS.

Innovation and market characteristics. The characteristics of the innovation and the market are discussed together since they are highly interdependent and even inseparable. Innovations can redefine existing markets, change the market structure, or even create totally new markets. On the other hand, the value of innovations is determined to a great extent by the characteristics of the market, such as demand and competition.

Theoretical economic literature suggests that the size of an innovation can have an effect on the propensity to patent the innovation. Denicolò and Franzoni (2003) assess the impact of the size of innovations on the propensity to patent in the context of the contract theory of patents and find that under the assumption of a linear demand function, innovations are more likely to be patented if they are large. This is because the rival has a greater incentive to duplicate the innovation if it is large, while patenting can be used to block duplication and secure monopoly profit for the duration of the patent. Horstmann et al. (1985) arrive at a similar conclusion when studying patents as information transfer mechanisms. They model a game of strategic patenting in which the rival can draw inferences about the innovator's private information on the basis of the patenting decision. Their reasoning for the finding is, however, very different from that of Denicolò and Franzoni (2003). Horstmann et al. (1985) argue that, in the context of a cost-reducing innovation, a greater cost reduction raises the innovator's output in the product market and thus makes imitation less attractive. Hence the decision to patent need not convey such a strong signal of unprofitability of imitation and patenting can be allowed to occur more often. Anton and Yao (2004), on the other hand, arrive at the opposite conclusion on the basis of their model of cost-reducing innovation. In the Anton and Yao model, patents offer limited protection while entailing disclosure of enabling knowledge to rivals as well as providing a signal of the total knowledge of the innovator. Anton and Yao (Ibid:3) argue that "… weak property rights imply disclosure incentives that are relatively stronger for smaller innovations, and as a result, larger innovations are protected more through secrecy as a response to the problem of imitation".

Protection from imitation—rather than signaling of cost-efficiency to competitors, which plays a central role in the Anton and Yao (2004) model—is constantly reported as the primary motive for patenting in innovation surveys (e.g. Duguet and Kabla, 1998; Cohen et al., 2000; Blind et al., 2006). Hence the hypothesis about the relationship between the size of innovations and the propensity to patent is based on the findings of Denicolò and Franzoni (2003) and Horstmann et al. (1985), which are also in line with the empirical investigations of De Melto et al. (1980) and Van der Panne and Kleinknecht (2005). This expectation is further buttressed when the assumption of the theoretical models that all innovations are patentable is relaxed. In order to be patentable, an invention has to be industrially applicable and of patentable subject matter, and it needs to satisfy the requirements of novelty and non-obviousness. Consequently, firms are likely to expect that patents be granted for large innovations with a higher probability than for smaller ones. This is probably taken into account when making the patenting decision. On the basis of these considerations, the following hypothesis is put forth:

Hypothesis 2: Large innovations are patented more frequently than smaller ones.

Measurement of the size of innovations or classification of innovations with respect to their size is a problematic issue even from the theoretical perspective. The complex and multidimensional nature of technological change makes it difficult to distinguish between large and small innovations, especially as innovations can be large in some dimensions while being small in others (cf. Henderson, 1993). The size—or radicalness—of an innovation can be defined, for instance, in terms of the technological novelty or magnitude of improvement and the socio-economic impact of the innovation (e.g. Schumpeter, 1912; Freeman and Perez, 1988), the magnitude of cost reduction and the economic implications of the innovation on the market structure (e.g. Arrow, 1962), or the effect the innovation has on the competencies of firms (e.g. Abernathy and Clark, 1985).

Furthermore, even if a certain theoretical definition of the size of an innovation is adopted, empirical measurement of the size is hardly straightforward. In order to address Hypothesis 2, the present study seeks to measure the size of the innovations by introducing four binary variables that capture different dimensions of the novelty and significance of the innovations. The variable NOVFIRM is coded as one for innovations that were specified as entirely new rather than major or minor improvements relative to the innovating firm's existing product by the survey respondent from the firm. Similarly, NOVMARK is coded as one if the innovation was specified to be new on the world market rather than just on the Finnish market. The variable SCIENCE seeks to proxy the technological novelty of the innovation. SCIENCE is coded as one if a new scientific breakthrough was specified as an important or very important (on a four-point Likert scale) factor for initiating the development of the innovation. Finally, the variable SIGNIF is introduced to pick out the truly significant innovations. This variable is based on a survey of experts drawn from industry, academia, and the public sector (see Hyvönen, 2001 for details of the survey and the data). The experts were asked to evaluate the significance[8] of the Sfinno innovations relating to their area of expertise on a four-point Likert scale (1–4). SIGNIF is coded as one if the mean score for the innovation is 3.5 or more.

Another attribute of innovations that can affect the propensity to patent is the complexity. Scherer (1983) and Levin et al. (1987), for instance, suggest that patenting of complex technological systems is more difficult than patenting of more discrete innovations. Levin et al. (1987) argue that the novelty of a discrete innovation can be relatively easily demonstrated in a patent application and infringement is relatively easy to verify when innovations are discrete. This is clearly more difficult to do for complex systems. Moreover, technological complexity can make innovations more difficult to imitate, thus reducing the need for patent protection. These arguments give rise to the following hypothesis:

Hypothesis 3a: Very complex innovations are patented less often than others.

The reasoning that led to Hypothesis 3a drew upon the impact of the technological and physical character of an innovation on the effectiveness and attractiveness of patents as a means for appropriation. On the other hand, complex technologies that are developed cumulatively may be subject to a high-degree of technological interdependence between competing firms (e.g. Cohen et al., 2000; Hall and Ziedonis, 2001). In such environments, firms can be highly dependent on cross-licensing as the intellectual property rights required to market a certain product get fragmented to a number of players. This is because such technological environments give rise to what Shapiro (2000:1–2) calls a patent thicket—that is, “a dense web of overlapping intellectual property rights that a company must hack its way through in order to actually commercialize new technology”. Firms may enter into patent portfolio races in order to improve their bargaining positions relative to others, leading them to patent inventions that would otherwise be left unpatented (Hall and Ziedonis, 2001). Hence the following hypothesis is suggested:

Hypothesis 3b: Cumulative technologies entail high propensities to patent.

In order to disentangle the different complexity-related effects proposed in Hypotheses 3a and 3b, two binary variables are constructed. First, the variable COMPLEX is designed to capture the technological and physical complexity of the innovations relevant for testing Hypothesis 3a. COMPLEX is coded as one if the innovation was classified as highly complex in Hyvönen’s 4-category taxonomy (e.g. Tanayama, 2002:56–57; Saarinen, 2005:160–161) by the VTT researchers. Hyvönen’s definition of a highly complex innovation is identical to the corresponding definition by Kleinknecht et al. (1993:44). Highly complex innovations are defined as systems consisting of numerous parts or components originating from different disciplines. Second, the variable CUMULTECH is designed to proxy the technological interdependence resulting from fragmentation of intellectual property rights (IPR) to cumulatively developing technologies (cf. Hypothesis 3b). CUMULTECH is coded as one if availability of a license was specified as an important or very important (on a four-point Likert scale) factor for initiating the development of the innovation.

One of the most robust findings emerging from the empirical literature is that the propensity to patent varies across industrial sectors. The origins of such differences are not entirely clear, however, since the variations can arise, for instance, as a result of the technological nature of the innovations or the characteristics of the markets. The software industry, for instance, probably experiences low propensities to patent because of issues related to the patentability of software rather than because of other attributes of the industry such as concentration. On the other hand, Denicolò and Franzoni (2003) argue that tight competition in the product market discourages duplication by the rival and thus makes patenting less attractive relative to secrecy for the innovator. Hence, while acknowledging the importance of controlling for differences in the technological nature of innovations in the empirical analysis, the following hypothesis is put forth:

Hypothesis 4: The propensity to patent declines with product market competition.

Unfortunately, empirical measurement of the degree of competition is a prevailing challenge in empirical industrial organization. This makes testing of Hypothesis 4 problematic. Measures of market concentration such as the Herfindahl–Hirschman index and concentration ratios follow standard definitions and can be objectively measured once the markets of interest are identified. However, such data is usually only readily available for industrial sectors and on a given level of aggregation and thus does not necessarily correspond to the relevant markets of interest. Consequently, a rough proxy emerging from the Sfinno data is used to measure the degree of competition in this study, instead of measures such as concentration ratios[9]. The binary variable PRICOMP is coded as one if price competition was specified as an important or very important (on a four-point Likert scale) factor for initiating the development of the innovation by the survey respondent from the firm. The usefulness of this variable as a proxy for the degree of product market competition hinges on the assumption that the ex post product market competition—that is, competition after the innovation is introduced to the market—correlates strongly enough with the ex ante competition—that is, competition before the market introduction of the innovation.

Control variables. In addition to the potential determinants of the propensity to patent outlined in the preceding hypotheses, several other factors can be expected to influence the propensity to patent. In order not to introduce omitted-variable bias into the estimates, it is imperative to control for such factors in the empirical analysis. The control variables used in this study are outlined below.

First, the results of Brouwer and Kleinknecht (1999), Van der Panne and Kleinknecht (2005), and Peeters and Van Pottelsberghe de la Potterie (2006) indicate that firms that engage in R&D collaboration exhibit higher propensities to patent than others. It is argued that this is due to the need to protect proprietary knowledge in the face of collaborative knowledge sharing and to clarify issues of ownership over co-developed innovations (e.g. Brouwer and Kleinknecht, 1999; Peeters and Van Pottelsberghe de la Potterie, 2006). The answers to the Sfinno survey question on whether the development of the innovation had involved collaboration with external partners give rise to a binary variable COLLAB that can be used to control for such an effect.

Second, it has been suggested that exporting activities tend to have a positive effect on the propensity to patent (e.g. Arundel and Kabla, 1998; Licht and Zoz, 1998). The Sfinno database contains information on whether the innovation has been exported. On the basis of this information, a binary variable INNOEXP is constructed.

Third, sets of dummy variables are introduced to control for differences in the propensity to patent across technology classes and time periods. Ten technology class dummies are constructed on the basis of the technology classification presented in Appendix 2. The dummies refer to the one-digit technology classes with the exception that the two-digit classes of 'agrochemistry and foodchemistry' and 'environmental technology' are picked out from their respective one-digit classes because the propensity to patent in these two-digit classes differs significantly from the propensity to patent in the rest of the one-digit class. Moreover, a set of eleven time period dummies is constructed so that for the early years as well as for the most recent years the time period classes contain more than one year. Such classification is used in order to have a sufficient number of observations in each time period class since the observations are not uniformly distributed in time, as shown in Figure 3. Appendix 1 summarizes all the variables introduced above.

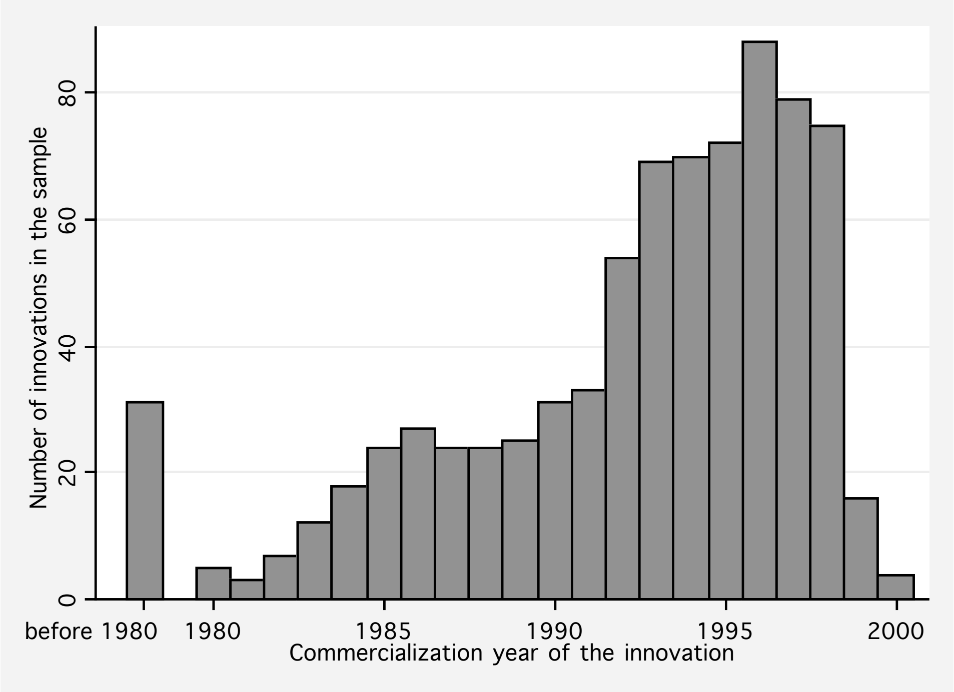

Figure 3. Distribution of observations over time.

4. Econometric analysis

This section lays out the econometric model to be estimated (Subsection 4.1) and presents the estimation results (Subsection 4.2).

4.1 Innovation-level model for the propensity to patent

Formulation of a model for the propensity to patent at the level of innovations requires an innovation-level definition of the propensity to patent. Following the frequency interpretation of probability associated with probability theorists such as John Venn (1876), the probability of an event can be interpreted as the relative frequency of occurrences of the event within a reference class. Hence the definition of the propensity to patent as 'the fraction of innovations for which at least one patent application is filed' gives rise to a corresponding probability interpretation. The propensity to patent can be understood as the probability that at least one patent application is filed for an innovation belonging to a given reference class (cf. Arora et al., 2003:6). More formally, the propensity to patent an innovation can be defined as the conditional probability:

$$\Pr\left[y=1\middle|\mathbf{x}\right], \text{ where } y=\begin{cases}1 & \text{if at least one patent application is filed, and}\\ 0 & \text{otherwise,}\end{cases} \tag{1}$$

and $\mathbf{x} \equiv (x_1, x_2, ..., x_K)$ is a vector of K variables that determines the reference class.

The probability definition of the propensity to patent allows for a formulation of a model for the propensity to patent in the spirit of random utility models (RUMs). Following Train (2003:18–21), the model is specified as follows:

i. An innovating firm files a patent application for its innovation if the (expected) payoff given the patent application, U_1, is higher than the (expected) payoff when no patent application is filed, U_0.

ii. U_1 and U_0 are known to the innovating firm, but not to the researcher. Instead, the researcher observes **x**, a vector of observable attributes of the innovation, the market, and the innovating firm.

iii. Following the random utility formulation, the payoffs are decomposed as

$$\begin{aligned} U_1 &= V_1(\mathbf{x}) + \varepsilon_1, \\ U_0 &= V_0(\mathbf{x}) + \varepsilon_0, \end{aligned} \tag{2}$$

where $V_1(\mathbf{x})$ and $V_0(\mathbf{x})$ are functions that relate the observed attributes, **x**, to the payoffs U_1 and U_0, respectively, and ε_1 and ε_0 capture the differences between U_1 and $V_1(\mathbf{x})$, and U_0 and $V_0(\mathbf{x})$, respectively. Because ε_1 and ε_0 are not known to the researcher, they are treated as random variables.

iv. The propensity to patent conditional on the observable attributes, **x**, can now be specified as

$$\begin{aligned}\Pr\left[y = 1 \middle| \mathbf{x}\right] &= \Pr\left[U_1 > U_0\right] \\ &= \Pr\left[V_1(\mathbf{x}) + \varepsilon_1 > V_0(\mathbf{x}) + \varepsilon_0\right] \\ &= \Pr\left[\varepsilon_0 - \varepsilon_1 < V_1(\mathbf{x}) - V_0(\mathbf{x})\right] \\ &= F\left(V_1(\mathbf{x}) - V_0(\mathbf{x})\right),\end{aligned} \tag{3}$$

where F is the cumulative distribution function of $\varepsilon \equiv \varepsilon_0 - \varepsilon_1$.

Following the conventional practice, $V_1(\mathbf{x})$ and $V_0(\mathbf{x})$ are assumed to be linear in parameters—that is, $V_1(\mathbf{x}) = \mathbf{x}'\boldsymbol{\beta}_1$ and $V_0(\mathbf{x}) = \mathbf{x}'\boldsymbol{\beta}_0$. Moreover, ε_0 and ε_1 are assumed to be distributed independently of **x**, while a natural behavioral assumption for ε_1 and ε_0 is that they are normally distributed. Hence, $\varepsilon \equiv \varepsilon_0 - \varepsilon_1$ is also normally distributed. Furthermore, an innocent normalization of the mean of ε to zero and the variance to unity is now possible as long as the model contains a constant term. Under these assumptions, the model for the propensity to patent becomes the standard probit model for binary choice:

$$\begin{aligned}\Pr\left[y = 1 \middle| \mathbf{x}\right] &= F\left(V_1(\mathbf{x}) - V_0(\mathbf{x})\right) \\ &= F\left(\mathbf{x}'(\boldsymbol{\beta}_1 - \boldsymbol{\beta}_0)\right) \\ &= F\left(\mathbf{x}'\boldsymbol{\beta}\right) \\ &= \Phi\left(\mathbf{x}'\boldsymbol{\beta}\right),\end{aligned} \tag{4}$$

where Φ is the standard normal cumulative distribution function and $\boldsymbol{\beta} \equiv \boldsymbol{\beta}_1 - \boldsymbol{\beta}_0$ is the vector of parameters to be estimated.

Owing to the object-based method of data collection, the Sfinno data contains multiple innovations from certain firms (cf. Figure 2); thus the observations are potentially subject to within-firm correlation due to unobserved firm-specific effects. Hence, the standard assumption of independence of observations fails, and the cluster sample characteristics of the data must be accounted for when the model is estimated.

In what follows, as in Wooldridge (2002), i indexes the cluster (i.e. the firm), g indexes the unit (i.e. the innovation), and N is the total number of firms and G_i the total number of innovations by firm i in the data. As the standard assumption of independence of observations fails, the specification of the joint distribution of $\mathbf{y}_i \equiv (y_{i1},\dots,y_{iG_i})$ conditional on $\mathbf{x}_i \equiv (\mathbf{x}_{i1},\dots,\mathbf{x}_{iG_i})$ for each cluster i becomes complicated. Hence the traditional maximum likelihood estimator (MLE) based on specification of $f(\mathbf{y} \mid \mathbf{x};\boldsymbol{\beta})$, the full joint density of **y** given **x**, cannot be readily utilized. However, assuming that the univariate densities $f_g(y_g \mid \mathbf{x}_g;\boldsymbol{\beta})$ are correctly specified for each g, the pooled probit model

$$\Pr\left[y_{ig} = 1 \middle| \mathbf{x}_{ig}\right] = \Phi\left(\mathbf{x}_{ig}'\boldsymbol{\beta}\right), \qquad g = 1,\dots,G_i \tag{5}$$

can be consistently estimated by a quasi-MLE that solves

$$\max_{\boldsymbol{\beta}} \sum_{i=1}^{N} \sum_{g=1}^{G_i} \log f_g\left(y_{ig} \middle| \mathbf{x}_{ig};\boldsymbol{\beta}\right), \tag{6}$$

where $\log f_g\left(y_{ig} \middle| \mathbf{x}_{ig};\boldsymbol{\beta}\right) = y_{ig} \log \Phi\left(\mathbf{x}'_{ig}\boldsymbol{\beta}\right) + \left(1 - y_{ig}\right) \log\left[1 - \Phi\left(\mathbf{x}'_{ig}\boldsymbol{\beta}\right)\right]$.

Wooldridge (2002) calls this the partial maximum likelihood estimator (PMLE). Consistency of the PMLE does not require that $\Pi_g f_g(y_g \mid \mathbf{x}_g;\boldsymbol{\beta})$ is the density of $\mathbf{y}$ given some set of conditioning variables. However, dependence of $y_1,\ldots,y_{Gi}$ results in the failure of the information matrix equality; thus cluster-robust asymptotic variance matrix and cluster-robust test statistics need to be computed instead of the usual ones. (See Wooldridge, 2002:401–410 for details on estimation and inference using the PMLE.)

The parameter estimates that result from the estimation of the pooled model are generally referred to as population averaged since the random effects are averaged out (Cameron and Trivedi, 2005:787). The population-averaged parameters should be expected to differ from those of an unobserved effects model which conditions also on the unobserved cluster-specific effects. However, the partial effects of the pooled model can be interpreted as the average partial effects (APEs)—that is, as partial effects averaged across the population distribution of the unobserved heterogeneity—of the unobserved effects model as long as the cluster-specific effects are independent of the included explanatory variables (see Wooldridge, 2002:22-24, 470–472, 482–490).

4.2 Estimation results

Table 1 contains the partial maximum likelihood estimates for two different specifications of the pooled probit model. The first specification of the pooled probit model (Pooled Probit 1) contains only the firm size dummies and the control variables. The purpose of this endeavor is to check whether the findings emerging from the Sfinno sample are consistent with the previous firm-level studies if the newly introduced variables are ignored. Moreover, estimation of this specification provides a point of reference for examining how the results change when the hypothesized determinants of the propensity to patent are accounted for. The second specification (Pooled Probit 2) contains all the variables introduced in Subsection 3.2 with the purpose of shedding light on the hypotheses outlined in that subsection.

Table 1. Estimation results for Pooled Probit 1 and 2.

Dependent variable: PATAPP		Pooled Probit 1			Pooled Probit 2	
Independent variables	Coef.	Robust Std. Err.	Partial effect◊	Coef.	Robust Std. Err.	Partial effect◊
Firm size classes						
EMP2	-0.3193**	0.1346	-0.1238**	-0.0341	0.1601	-0.0136
EMP3	-0.5223***	0.1471	-0.2046***	-0.2512	0.1804	-0.0993
EMP4	-0.2690	0.2465	-0.1038	-0.2605	0.2575	-0.1028
Other firm characteristics						
PATENTS				0.0160**	0.0070	0.0064**
STARTUP				0.3038**	0.1407	0.1206**
Innovation and market characteristics						
SIGNIF				0.6566**	0.2974	0.2463**
NOVFIRM				0.4636***	0.1306	0.1831***
NOVMARK				0.8850***	0.1148	0.3362***
SCIENCE				0.3118**	0.1483	0.1231**
COMPLEX				-0.5699**	0.2740	-0.2170**
CUMULTECH				0.5651**	0.2281	0.2155***
PRICOMP				-0.2415*	0.1294	-0.0960*
Technology classes (ref. CONSUM)						
ELECTRO	-0.1712	0.3063	-0.0653	-0.1031	0.3329	-0.0411
INSTRU	-0.0997	0.2887	-0.0376	-0.0642	0.3161	-0.0256
CHEM	0.2382	0.3723	0.0837	0.1622	0.4053	0.0644
AGRI&FOODCHEM	-0.4295	0.3431	-0.1677	-0.2988	0.3592	-0.1176
PROCTECH	0.0990	0.2891	0.0360	0.3322	0.3119	0.1300
ENVIRO	1.1122**	0.4614	0.2836**	1.1652**	0.5231	0.3767**
MACH	0.2091	0.2871	0.0740	0.5288*	0.3201	0.2011
EARTH&WATER	0.0601	0.3403	0.0220	0.0773	0.3626	0.0308
SOFT	-1.4929***	0.3834	-0.5155***	-1.4103***	0.4025	-0.4225***
Time periods (10 dummies)		See Appendix 3 for the estimates			See Appendix 3 for the estimates	
Other control variables						
COLLAB	0.1903	0.1416	0.0722	-0.0717	0.1430	-0.0286
INNOEXP	0.2360**	0.1150	0.0885**	0.0552	0.1186	0.0220
Constant	0.0626	0.3204		-1.0853***	0.3652	
Robust Wald tests for joint hypotheses		χ^2 (df)	p-value		χ^2 (df)	p-value
H_0: All coefs zero (exc. constant)		82.21 (24)	0.0000		189.51 (33)	0.0000
H_0: All firm size class coefs zero		13.67 (3)	0.0034		2.60 (3)	0.4570
H_0: All tech. class coefs zero		47.95 (9)	0.0000		50.60 (9)	0.0000
H_0: All time period coefs zero		16.56 (10)	0.0846		22.78 (10)	0.0116
Number of observations		791			791	
Number of clusters		555			555	
Log pseudolikelihood		-461.3642			-393.72924	
McFadden's pseudo R^2		0.145			0.270	
Efron's pseudo R^2		0.187			0.335	
McKelvey and Zavoina's pseudo R^2		0.288			0.472	
Percent correctly predicted						
for observations with PATAPP=1		88.77			86.12	
for observations with PATAPP=0		41.84			66.77	
for all observations		68.77			77.88	

Significance level notation: *** 1%, ** 5%, * 10%.

◊The partial effects are estimated at a point where firm size, technology class, and time period dummies are all zero and other variables are assigned their mean values. The partial effects are computed as discrete changes in the propensity to patent for binary variables and as a partial derivative for the variable PATENTS. The significance level notation for the partial effects is based on standard errors computed using the delta method.

Hypotheses 1a and 1b. The estimation results for Pooled Probit 1 show a non-monotonic U-shaped relationship between firm size and the propensity to patent. This finding can be argued to be in accordance with the survey evidence of the positive relationship between firm size and the propensity to patent (e.g. Arundel and Kabla, 1998; Duguet and Kabla, 1998; Arora et al., 2003) since the firm-level surveys have largely ignored the smallest firms. The results for Pooled Probit 1 suggest that among the relatively large firms, the propensity to patent increases with firm size. While being ignored in the firm-level studies, small start-up ventures are well represented in the Sfinno sample. Pooled Probit 2 provides statistically significant (at the 5 percent significance level) evidence of relatively high propensities to patent in start-up ventures, thus lending support to Hypothesis 1b. Moreover, Pooled Probit 2 lends support to Hypothesis 1a, which holds that the propensity to patent increases with the scale of patenting, by showing a positive and statistically significant (at the 5 percent significance level) effect of the variable PATENTS on the propensity to patent. The U-shaped relationship between firm size and the propensity to patent appears to be captured relatively well by the variables for start-up ventures (STARTUP) and the scale of patenting (PATENTS). Once STARTUP and PATENTS are included in the model, the null hypothesis of the coefficients of the firm size dummies all being zero can no longer be rejected at any meaningful level of significance. Unfortunately, as discussed in Subsection 3.2, the variable designed to account for the scale of patenting may be subject to simultaneous causality. Such endogeneity can compromise the validity of the evidence in support of Hypothesis 1a. However, it is somewhat reassuring that if the variable PATENTS is excluded from the model, the coefficient of EMP4 increases as expected. The exogeneity assumption of PATENTS will be formally tested at the end of this subsection.

Hypothesis 2. The results for Pooled Probit 2 provide support to Hypothesis 2, which proposes that large innovations are patented more often than others. All variables designed to capture different dimensions of the size of innovations (SIGNIF, NOVFIRM, NOVMARK, SCIENCE) display positive coefficients and sizeable positive partial effects. The coefficients and partial effects can be concluded to differ from zero at least at the 5 percent significance level.

Hypotheses 3a and 3b. The estimation results lend support to the hypotheses related to the effect of the complexity of innovations on the propensity to patent. The variables designed to capture the technological complexity of the innovations (COMPLEX) and the fragmentation of intellectual property rights (IPR) to cumulatively developing technology (CUMULTECH) help to disentangle the opposite complexity-related effects discussed in Subsection 3.2. First, the coefficient and partial effect of COMPLEX are negative and statistically different from zero (at the 5 percent significance level), suggesting that very complex innovations are patented less often than others—as proposed in Hypothesis 3a. Second, the coefficient and partial effect of CUMULTECH are positive and statistically different from zero (at least at the 5 percent significance level). The finding that dependence on the availability of a license in the development of an innovation increases the propensity to patent indicates that fragmentation of IPR encourages patenting and supports the proposition of Hypothesis 3b that cumulative technologies entail high propensities to patent.

Hypothesis 4. Pooled Probit 2 shows a negative coefficient for the variable designed as a proxy for the degree of competition in the product market. The coefficient and partial effect appear to differ from zero only at the 10 percent significance level, lending limited support to Hypothesis 4, which proposes that the propensity to patent declines with competition in the product market. Moreover, this result needs to be taken with a grain of salt since price competition in the product market might be expected to trigger product differentiation and incremental change rather than development of large innovations (cf. Tanayama, 2002). If the variables designed to measure the size of innovations fail to capture the effect of the size on the propensity to patent in its entirety, it is possible that price competition is negatively associated with the propensity to patent because it affects the type of innovative activity rather than the propensity to patent directly.

Control variables. The results of Table 1 lend significant support to the assumption that the propensity to patent varies across technologies. As expected, there seems to be a relatively high tendency to patent machinery (MACH) and chemicals and pharmaceuticals (CHEM), and a relatively low propensity to patent software (SOFT). Interestingly, environmental technology (ENVIRO) seems to experience a very high patenting propensity. This may well be because the rising concerns about sustainable development and global warming are making environmental technology increasingly important, and the early innovators in this growing field might seek to secure a share of returns to the later-generation innovations in the course of cumulative development of the technology in the future.

The null hypothesis that the coefficients of all the time period dummies are zero can be rejected in Pooled Probit 1 and 2 at the 10 and 5 percent significance levels, respectively. The estimation results (see Appendix 3) clearly provide no evidence of a general increase in the propensity to patent significant product innovations that would explain the recent surge in patenting (cf. Kortum and Lerner, 1999; Hall and Ziedonis, 2001; and Hall, 2005).

The need to protect proprietary knowledge in the face of collaborative knowledge sharing and to clarify issues of ownership over co-developed innovations has been argued to increase the propensity to patent in firms that engage in R&D collaboration. Peeters and Van Pottelsberghe de la Potterie (2006) refer to this as the ‘need’ effect of R&D collaboration on the propensity to patent. In order to control for such an effect, COLLAB appears as a control variable in both specifications of Table 1. However, the estimation results provide no significant evidence of such a relationship between R&D collaboration and the propensity to patent. This indicates that the finding of a positive relationship between R&D collaboration and the propensity to patent in firm-level studies such as Brouwer and Kleinknecht (1999) might be due to what Peeters and Van Pottelsberghe de la Potterie (2006) call the ‘novelty’ effect—that is, the tendency of R&D collaboration to lead to the generation of more ‘fundamental and breakthrough knowledge’ than in-house R&D[10]. Since the present study seeks to control for the effect of the size of innovations on the propensity to patent directly, the 'novelty' effect should be captured by the innovation-size variables rather than the R&D collaboration variable.

Similarly, the positive relationship between exporting activities and the propensity to patent observed, for instance, in Licht and Zoz (1998) and Arundel and Kabla (1998) may result from exporting firms developing larger innovations—or better yet, firm's with larger innovations choosing to export them—rather than having an inherently higher propensity to patent. Such an argument is supported by the observation that in Pooled Probit 1 the control variable INNOEXP appears to have a positive and statistically significant (at the 5 percent significance level) effect on the propensity to patent, but once the size of innovations is controlled for, evidence of such an effect no longer exists (cf. Pooled Probit 2).

Testing for the exogeneity assumption of PATENTS. The exogeneity assumption of PATENTS is tested by a two-step procedure in the spirit of Smith and Blundell (1986) and Rivers and Vuong (1988). In practice, the test can be applied as follows (see Wooldridge, 2002:472–478 for details). First, the potentially endogenous variable is regressed (using the standard OLS method) on the exogenous variables of the probit model and at least one additional instrument. Second, the probit model is estimated with the exogenous variables, the potentially endogenous variable, and the residuals of the first-stage regression as explanatory variables. Then the test of the null hypothesis of exogeneity can be based on the significance of the residual in the second-stage probit. Since the distribution of the first-stage error term plays no role under the null, such a test is valid without assuming normality or homoscedasticity of the first-stage error term and the test can be applied very broadly, even if the potentially endogenous variable is not continuous (Wooldridge, 2002:474).

Identification of the second-stage probit requires that at least one of the explanatory variables of the first-stage regression be excluded from the probit model. The firm size dummies are natural candidates for instruments to be excluded from the probit model since they are important determinants of the scale of patenting but are not expected to affect the propensity to patent directly. The size-related hypotheses of Subsection 3.2 propose that the start-up status and the scale of patenting are responsible for the association between size and scale and the propensity to patent, while the null hypothesis that the firm size dummies can be excluded from the innovation-level model for the propensity to patent cannot be rejected once these factors are controlled for (cf. Table 1). Furthermore, firm size should not be subject to simultaneous causality that threatens the patenting-scale variable since the decision of whether or not to patent an innovation hardly affects the size of the innovating firm—at least in the short run.

Table 2 presents the results for the test of the exogeneity assumption of PATENTS. The test results indicate that the null hypothesis of exogeneity of PATENTS cannot be rejected at meaningful levels of significance. This supports the validity of PATENTS as a measure of the scale of patenting in the model. The validity of the test naturally hinges on the assumption that the instruments for the potentially endogenous variable are themselves exogenous.

Table 2: Testing the exogeneity assumption of PATENTS.

Dependent variable in the probit model: PATAPP	
Explanatory variables in the probit model	
Potentially endogenous variable	PATENTS
Exogenous variables	STARTUP
	SIGNIF
	NOVFIRM
	NOVMARK
	SCIENCE
	COMPLEX
	CUMULTECH
	PRICOMP
	Technology class dummies (9)
	Time period dummies (10)
	COLLAB
	INNOEXP
Instruments	Firm size dummies (3)
Test of exogeneity of PATENTS	
H_0: Coef of the OLS residual zero in the probit model	
Robust asymptotic t-statistic	0.89
p-value	0.373

5. Conclusion

Thus far most of the empirical investigations into the propensity to patent have been confined to the use of industry and firm-level data, and the failure to control for innovation-level factors has made the interpretation of the results somewhat problematic. The observed variations in the propensity to patent across industries and firms might reflect differences in the characteristics of innovations developed in these industries and firms rather than some inherent differences in the propensity to patent. Moreover, the absence of innovation-level variables has rendered innovation-related hypotheses emerging from the theoretical literature untestable in the industry and firm-level studies. This study seeks to shed new light on the propensity to patent at the innovation level, while also contributing to the long tradition of research on the relationship between firm size and the propensity to patent. By taking the analysis to the innovation level, this study also brings the empirics closer to the theoretical work on the propensity to patent.

The present study set out to cast new light on the question of how the propensity to patent an innovation is affected by the characteristics of the innovation, the market, and the innovating firm. The innovation-level model for the propensity to patent was derived in the spirit of random utility models, and the emerging probit model was estimated on a sample of 791 Finnish product innovations drawn from the Sfinno database compiled at VTT Technical Research Centre of Finland.

The results from the econometric analysis indicate that various characteristics of the innovation, the market, and the innovating firm have a significant effect on the propensity to patent. First, there appears to be a U-shaped relationship between firm size and the propensity to patent, which can be attributed to a relatively large extent to economies of scale in the patenting activity as well as to the relatively important role of patenting in start-up ventures. Second, the estimation results suggest that larger—that is, more novel and significant—innovations are patented more frequently than smaller ones. Third, technologically very complex innovations appear to be patented less often than others, while the fragmentation of intellectual property rights to cumulatively developing technology seems to entail high propensities to patent. Fourth, the econometric analysis produces weak evidence on a negative relationship between the propensity to patent and the product market competition. This evidence needs to be taken with a grain of salt, however, since intense price competition in the product market might indirectly affect the propensity to patent by affecting the size of the innovations rather than by having a direct impact on the propensity to patent. Furthermore, certain factors—such as R&D collaboration and exporting activities—that have appeared to have an impact on the propensity to patent in the firm-level studies fail to exhibit a statistically significant effect once the innovation-level factors are controlled for. This might be indicative of such variables having only an indirect effect since they may well be associated with the size of innovations rather than affecting the propensity to patent directly. While this study seeks to capture different dimensions of the size of innovations with some success using a number of qualitative variables, development of more accurate measures of the size of innovations should make it easier to disentangle the direct effects from the indirect effects that influence patenting through the size of innovations.

The results outlined above should be of obvious interest to those who depend on patent data in drawing conclusions about innovation and technological change. The finding that larger product innovations are patented more frequently than smaller ones should be comforting news from the perspective of using patents as an economic indicator of innovation since it implies that large innovations enter the patent indicator at a relatively high probability. However, the study also points to the weaknesses of patent data by demonstrating that the propensity to patent varies significantly across firms and technologies. For instance, the evidence in favor of the hypotheses proposing that the propensity to patent increases with the scale of patenting and that start-up ventures exhibit high propensities to patent suggests that patents are a rather problematic measure of innovations in the context of testing the Schumpeterian hypotheses.

Moreover, the size-related hypotheses suggest that small start-up ventures are more dependent on patent protection than larger firms while experiencing a disadvantage in obtaining and enforcing patents. This should have important implications for the optimal design of the patent system since it is highly probable that not all valuable ideas originate in the large corporations and thus also small entities need to be provided with sufficient incentives for developing their ideas into innovations. Harnessing the innovative capacity of small firms is clearly an important challenge for any economy.

Because in reality an innovation can be protected by a number of patents, a single patent can cover numerous innovations, and not all patents relate to innovations, a complete investigation of the extent to which patents are representative of different innovations is beyond the scope of this study. Furthermore, the nature of the data used in this study does not allow for consideration of process innovations only developed for the firms' internal use. Clearly, further research is needed to paint a clear picture of the relationship between innovations and patents and to answer the question of the extent to which patents are representative of the wider universe of innovations. All in all, the study provides a rather encouraging perspective of the potential of innovation-level investigations in contributing to our understanding of the features and patterns of technological activities. This study is just a small step in trying to shed light on the complex relationship between patents and innovations that has been remained extremely elusive thus far. Nevertheless, the results indicate that this line of research can prove a very valuable complement to different industry and firm-level investigations.

NOTES

---

[1] Much of the theoretical work on patents leaves the decision to patent unmodeled and assumes that all (patentable) innovations are patented.

[2] Following the relevant theoretical literature (Denicolò and Franzoni, 2003; Anton and Yao, 2004), the term *size* (large vs. small) of an innovation is adopted in the present study instead of relatively synonymous alternatives such as the *radicalness* (radical vs. incremental) of an innovation.

[3] Two famous hypotheses associated with Schumpeter (1942) claim that (1) innovation increases more than proportionally with firm size and (2) there is a positive relationship between innovation and market concentration.

[4] This study follows the Sfinno-project in defining an innovation as an invention that has been commercialized on the market by a business firm or an equivalent (Palmberg et al., 1999:38, 2000:10; Saarinen, 2005:19–20).

[5] Figure 1 is a refined version of the figure in Basberg (1987:133).

[6] The results may also be biased due to the shortcomings of R&D expenditure as an indicator of innovation inputs; formal R&D is only one of the innovation inputs and standard innovation surveys tend to underestimate the R&D activities of small firms (e.g. Patel and Pavitt, 1995; Kleinknecht et al., 2002).

[7] Using the percentage of innovations, rather than inventions, overcomes the drawback—inherent in Mansfield's definition—that many inventions are never commercialized and hence have little economic significance. Moreover, the innovations of interest should not be limited to patentable innovations because the propensity to patent figures are of interest as an indicator of the extent to which patents represent the whole population of innovations. (Arundel and Kabla, 1998.)

[8] The definition of a significant innovation adopted for the survey is that the innovation has to be economically and technologically significant and apart from economic success may have had significant impact on the industry (Hyvönen, 2001:4).

[9] Concentration ratios (e.g. CR3, CR5, CR10) based on the NACE classification (General Industrial Classification of Economic Activities within the European Communities) at the three-digit level were also tested as measures of product market competition but they failed to be statistically significant in any of the specifications by a wide margin.

[10] Peeters and Van Pottelsberghe de la Potterie (2006) argue that the 'need' effect should dominate in collaboration arrangements with competitors, while the 'novelty' effect should dominate in partnerships with scientific institutions. Mäkinen (2007) further disaggregates COLLAB into collaboration with universities and research institutes, competitors, subcontractors, and customers, and finds that only collaboration with universities and research institutes has a statistically significant effect on the propensity to patent. This could be interpreted as evidence that it is the 'novelty' effect rather than the 'need' effect that drives the results in the firm-level studies.

References

Abernathy, W.J. and Clark, K.B. (1985). Innovation: Mapping the Winds of Creative Destruction. Research Policy, 14, 3–22.

Anton, J.J. and Yao, D.A. (2004). Little Patents and Big Secrets: Managing Intellectual Property. RAND Journal of Economics, 35, 1–22.

Archibugi, D. and Pianta, M. (1996). Measuring Technological Change through Patents and Innovation Surveys. Technovation, 16, 451–468.

Arora, A., Ceccagnoli, M. and Cohen, W.M. (2003). R&D and the Patent Premium. NBER Working Paper No. 9431.

Arora, A., Fosfuri, A. and Gambardella, A. (2001). Markets for Technology. Cambridge, MA: MIT Press.

Arrow, K. (1962). Economic Welfare and the Allocation of Resources for Invention. In R. Nelson (Ed.): The Rate and Direction of Inventive Activity: Economic and Social Factors. Princeton: Princeton University Press.

Arundel, A. and Kabla, I. (1998). What Percentage of Innovations Are Patented? Empirical Estimates for European Firms. Research Policy, 27, 127–141.

Barnett, J.M. (2003). Private Protection of Patentable Goods. Fordham University School of Law Research Paper 28.

Basberg, B.L. (1987). Patents and the Measurement of Technological Change: A Survey of the Literature. Research Policy, 16, 131–141.

Blind, K., Edler, J., Frietsch, R. and Schmoch, U. (2006). Motives to Patent: Empirical Evidence from Germany. Research Policy, 35, 655–672.

Bound, J., Cummins, C., Griliches, Z., Hall, B.H. and Jaffe, A. (1984). Who Does R&D and Who Patents? In Z. Griliches (Ed.): R&D, Patents, and Productivity. Chicago, IL: University of Chicago Press.

Brouwer, E. and Kleinknecht, A. (1999). Innovative Output and a Firm's Propensity to Patent: An Exploration of CIS Micro Data. Research Policy, 28, 615–624.

Cameron, A.C. and Trivedi, P.K. (2005). Microeconometrics: Methods and Applications. Cambridge: Cambridge University Press.

Cohen, W.M. and Klepper, S. (1996). A Reprise of Size and R&D. Economic Journal, 106, 925–951.

Cohen, W.M., Nelson, R.R. and Walsh, J.P. (2000). Protecting their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not). NBER Working Paper No. 7552.

De Melto, D.P., McMullen, K.E. and Wills, R.M. (1980). Preliminary Report: Innovation and Technological Change in Five Canadian Industries. Economic Council of Canada Discussion Paper No. 176.

Denicolò, V. and Franzoni, L.A. (2003). The Contract Theory of Patents. International Review of Law and Economics, 23, 365–380.

Duguet, E. and Kabla, I. (1998). Appropriation Strategy and the Motivations to use the Patent System: An Econometric Analysis at the Firm Level in French Manufacturing. Annales D'Économie et de Statistique, 49/50, 289–327.

Freeman, C. and Perez, C. (1988). Structural Crises of Adjustment: Business Cycles and Investment Behaviour. In G. Dosi, C. Freeman, R. Nelson, G. Silverberg, and L. Soete (Eds.): Technical Change and Economic Theory. London: Pinter Publishers.

Gallini, N.T. (1992). Patent Policy and Costly Imitation. RAND Journal of Economics, 23, 52–63.

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

Hall, B.H. (2005). Exploring the Patent Explosion. Journal of Technology Transfer, 30, 35–48.

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

Hall, B.H. and Ziedonis, R.H. (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.

Henderson, R. (1993). Underinvestment and Incompetence as Responses to Radical Innovation: Evidence from the Photolithographic Alignment Equipment Industry. RAND Journal of Economics, 24, 248–270.

Hortsmann, I., MacDonald, G.M. and Slivinski, A. (1985). Patents as Information Transfer Mechanisms: To Patent or (Maybe) Not to Patent. Journal of Political Economy, 93, 837–858.

Hyvönen, J. (2001). Significant Innovations in Finland. A Paper Prepared for the ETIC Session, Strasbourg, France, October 8–18, 2001.

Kleinknecht, A. and Bain, D. (Eds.) (1993). New Concepts in Innovation Output Measurement. Houndmills: Macmillan Press.

Kleinknecht, A., Reijnen, J.O.N. and Smits, W. (1993). Collecting Literature-based Innovation Output Indicators: The Experience in the Netherlands. In A. Kleinknecht and D. Bain (Eds.): New Concepts in Innovation Output Measurement. Houndmills: Macmillan Press.

Kleinknecht, A., Van Montfort, K. and Brouwer, E. (2002). The Non-trivial Choice between Innovation Indicators. Economics of Innovation and New Technology, 11, 109–121.

Kortum, S. and Lerner, J. (1999). What Is Behind the Recent Surge in Patenting? Research Policy, 28, 1–22.

Kortum, S. and Lerner, J. (2000). Assessing the Contribution of Venture Capital to Innovation. RAND Journal of Economics, 31, 674–692.

Kultti, K., Takalo, T. and Toikka, J. (2007). Secrecy vs. Patenting. Forthcoming in RAND Journal of Economics.

Lanjouw, J.O. and Schankerman, M. (2004). Protecting Intellectual Property Rights: Are Small Firms Handicapped? Journal of Law and Economics, 47, 45–74.

Leppälahti, A. (2000). Comparisons of Finnish Innovation Surveys. Statistics Finland; Science, Technology and Research 2000:1.

Lerner, J. (1995). Patenting in the Shadow of Competitors. Journal of Law and Economics, 38, 463–495.

Levin, R.C., Klevorick, A.K., Nelson, R.R., and Winter, S.G. (1987). Appropriating the Returns from Industrial R&D. Brookings Papers on Economic Activity, 3, 783–820.

Licht, G. and Zoz, K. (1998). Patents and R&D: An Econometric Investigation Using Applications for German, European and US Patents by German Companies. Annales D'Économie et de Statistique, 49/50, 329–360.

Long, C. (2002). Patent Signals. University of Chicago Law Review, 69, 625–679.

Mansfield, E. (1986). Patents and Innovation: An Empirical Study. Management Science, 32, 173–181.

Mäkinen, I. (2007). To Patent or Not to Patent? An Innovation-Level Investigation of the Propensity to Patent. VTT Publications 646.

OECD (1997). Proposed Guidelines for Collecting and Interpreting Technological Innovation Data: The Oslo Manual. Paris: OECD and Eurostat.

Palmberg, C. (2001). Sectoral Patterns of Innovation and Competence Requirements: A Closer Look at Low-Tech Industries. Sitra Reports Series No. 8.

Palmberg, C., Leppälahti, A., Lemola, T. and Toivanen, H. (1999). Towards a Better Understanding of Innovation and Industrial Renewal in Finland – a New Perspective. VTT Group for Technology Studies Working Paper No. 41/99.

Palmberg, C., Niininen, P., Toivanen, H. and Wahlberg, T. (2000) Industrial Innovation in Finland: First Results of the Sfinno-project. VTT Group for Technology Studies Working Paper No. 47/00.

Parchomovsky, G. and Wagner, R.P. (2005). Patent Portfolios. U of Penn. Law School Public Law Working Paper 56.

Patel, P. and Pavitt, K. (1995). Patterns of Technological Activity: their Measurement and Interpretation. In P. Stoneman (Ed.): Handbook of the Economics of Innovation and Technological Change. Oxford: Blackwell Publishers.

Peeters, C. and Van Pottelsberghe de la Potterie, B. (2006). Innovation Strategy and the Patenting Behavior of Firms. Journal of Evolutionary Economics, 16, 109–135.

Reinganum, J.F. (1983). Uncertain Innovation and the Persistence of Monopoly. American Economic Review, 73, 741–748.

Rivers, D. and Vuong, Q.H. (1988). Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models. Journal of Econometrics, 39, 347–366.

Saarinen, J. (2005). Innovations and Industrial Performance in Finland 1945–1998. Doctoral Dissertation, Lund University, Lund Studies in Economic History.

Scherer, F.M. (1965). Firm Size, Market Structure, Opportunity, and the Output of Patented Inventions. American Economic Review, 55, 1097–1125.

Scherer, F.M. (1983). The Propensity to Patent. International Journal of Industrial Organization, 1, 107–128.

Schmookler, J. (1966). Invention and Economic Growth. Cambridge, MA: Harvard University Press.

Schumpeter, J. (1912). Theorie der wirtschaftlichen Entwicklung. Leipzig: Duncker & Humboldt. English translation (1934). The Theory of Economic Development. Cambridge, MA: Harvard University Press.

Schumpeter, J. (1942). Capitalism, Socialism and Democracy. New York: Harper and Brothers.

Scotchmer, S. and Green, J. (1990). Novelty and Disclosure in Patent Law. RAND Journal of Economics, 21, 131–146.

Shapiro, C. (2000). Navigating the Patent Thicket: Cross Licenses, Patent Pools, and Standard-Setting. Competition Policy Center, University of California, Berkeley, Working Paper No. CPC00-11.

Smith, R.J. and Blundell, R.W. (1986). An Exogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labor Supply. Econometrica, 54, 679–686.

Takalo, T. (1998). Innovation and Imitation under Imperfect Patent Protection. Journal of Economics, 67, 229–241.

Tanayama, T. (2002). Empirical Analysis of Processes Underlying Various Technological Innovations. VTT Publications 463.

Taylor, C.T. and Silberston, Z.A. (1973). The Economic Impact of the Patent System. Cambridge: Cambridge University Press.

Train, K.E. (2003). Discrete Choice Methods with Simulation. Cambridge: Cambridge University Press.

Van der Panne, G. and Kleinknecht, A. (2005). The Propensity to Patent an Innovation. An Empirical Investigation. A Paper Presented at the Fourth European Meeting on Applied Evolutionary Economics (EMAEE), Utrecht, The Netherlands, 19–21 May 2005.

Venn, J. (1876). The Logic of Chance (2nd Ed.). London: Macmillan.

Waterson, M. (1990). The Economics of Product Patents. American Economic Review, 80, 860–869.

Wooldridge, J.M. (2002). Econometric Analysis of Cross Section and Panel Data. Cambridge, MA: MIT Press.

APPENDIX 1: Summary of the variables

Dependent variable	Definition	Type	Mean	St. Dev.
PATAPP	Patent application was filed for the innovation (yes/no)	1/0	0.5740	0.4948
Explanatory variables				
Firm characteristics				
EMP	Number of employees in the firm at the year of the commercialization	#	1113.873	2495.891
EMP1	0-9 employees in the firm at the year of the commercialization (yes/no)	1/0	0.3552	0.4789
EMP2	10-99 employees in the firm at the year of the commercialization (yes/no)	1/0	0.2149	0.4110
EMP3	100-999 employees in the firm at the year of the commercialization (yes/no)	1/0	0.2048	0.4038
EMP4	1000 or more employees in the firm at the year of the commercialization (yes/no)	1/0	0.2250	0.4179
STARTUP	The firm was defined as a start-up developing an innovation (yes/no)	1/0	0.3603	0.4804
PATENTS	Number of patent applications filed by the firm the year before the development of the innovation started	#	3.4083	11.5773
Innovation and market characteristics				
SIGNIF	The innovation was specified as very significant by experts (yes/no)	1/0	0.0518	0.2218
NOVFIRM	The innovation was entirely new to the firm (yes/no)	1/0	0.6157	0.4867
NOVMARK	The innovation was new to the world market (yes/no)	1/0	0.7206	0.4490
SCIENCE	Scientific breakthrough was important for initiating the development of the innovation (yes/no)	1/0	0.1555	0.3626
COMPLEX	The innovation was specified as very complex by experts (yes/no)	1/0	0.0291	0.1681
CUMULTECH	Availability of a license was important for initiating the development of the innovation (yes/no)	1/0	0.0582	0.2342
PRICOMP	Price competition was important for initiating the development of the innovation (yes/no)	1/0	0.2781	0.4484
Technology classes				
CONSUM	The innovation belongs to 1-digit technology class 60 'Consumption goods and equipment' (yes/no)	1/0	0.0329	0.1784
ELECTRO	The innovation belongs to 1-digit technology class 10 'Electrotechnology' (yes/no)	1/0	0.0860	0.2805
INSTRU	The innovation belongs to 1-digit technology class 20 'Instruments' (yes/no)	1/0	0.1416	0.3489
CHEM	The innovation belongs to 1-digit technology class 30 'Chemistry, pharmaceutical technology' excluding 35 (yes/no)	1/0	0.0594	0.2366
AGRI&FOODCHEM	The innovation belongs to 2-digit technology class 35 'Agrochemistry, foodchemistry' (yes/no)	1/0	0.0544	0.2269
PROCTECH	The innovation belongs to 1-digit technology class 40 'Process technology, special equipment' exluding 48 (yes/no)	1/0	0.2579	0.4378
ENVIRO	The innovation belongs to 2-digit technology class 48 'Environmental technology' (yes/no)	1/0	0.0253	0.1571
MACH	The innovation belongs to 1-digit technology class 50 'Mechanical engineering, equipment' (yes/no)	1/0	0.1884	0.3913
EARTH&WATER	The innovation belongs to 1-digit technology class 70 'Earth construction and hydraulic engineering, mining' (yes/no)	1/0	0.0367	0.1881
SOFT	The innovation belongs to 1-digit technology class 80 'Software' (yes/no)	1/0	0.1176	0.3223
Time periods				
PRE1986	The innovation was commercialized before 1986 (yes/no)	1/0	0.1264	0.3325
YEARS86-87	The innovation was commercialized in 1986-87 (yes/no)	1/0	0.0645	0.2458
YEARS88-89	The innovation was commercialized in 1988-89 (yes/no)	1/0	0.0619	0.2412
YEARS90-91	The innovation was commercialized in 1990-91 (yes/no)	1/0	0.0809	0.2729
YEAR1992	The innovation was commercialized in 1992 (yes/no)	1/0	0.0683	0.2524
YEAR1993	The innovation was commercialized in 1993 (yes/no)	1/0	0.0872	0.2824
YEAR1994	The innovation was commercialized in 1994 (yes/no)	1/0	0.0885	0.2842
YEAR1995	The innovation was commercialized in 1995 (yes/no)	1/0	0.0910	0.2878
YEAR1996	The innovation was commercialized in 1996 (yes/no)	1/0	0.1113	0.3146
YEAR1997	The innovation was commercialized in 1997 (yes/no)	1/0	0.0999	0.3000
POST1997	The innovation was commercialized after 1997 (yes/no)	1/0	0.1201	0.3253
Other control variables				
COLLAB	Collaboration was associated with the development of the innovation (yes/no)	1/0	0.8698	0.3368
INNOEXP	The innovation had been exported (yes/no)	1/0	0.6523	0.4765

APPENDIX 2: Technology classification, VTT Innovation Studies

Technology class	**IPC-class**
10 Electrotechnology	
11 Electrical machinery and equipment, electric energy	F21; G05F; H01B,C,F,G,H,J,K,M, R,T; H02; H05B,C,F,K
12 Audiovisual technology	G09F,G; G11B; H03F,G,J; H04N-003,-005,-009,-013,015,-017,R,S
13 Telecommunications	G08C; H01P,Q; H03B,C,D,H,K, L,M; H04B,H,J,K,L,M,N-001,-007,-011,Q
14 Information technology	G06; G11C; G10L
15 Semiconductors	H01L
20 Instruments	
21 Optics	G02; G03B,C,D,F,G,H; H01S
22 Analysis, measurement, and control technology	G01B,C,D,F,G,H,J,K,L,M,N,P,R,S,V,W; G04; G05B,D; G07; G08B,G; G09B,C,D; G12
23 Healthcare technology	A61B,C,D,F,G,H,J,L,M,N
24 Nuclear technology	G01T; G21; H05G,H
30 Chemistry, pharmaceutical technology	
31 Organic chemistry	C07C,D,F,H,J,K
32 Macromolecule chemistry, polymer chemistry	C08B,F,G,H,K,L; C09D,J; C13L
33 Pharmaceutical technology, cosmetics	A61K
34 Biotechnology	C07G; C12M,N,P,Q,R,S
35 Agrochemistry, foodchemistry	A01H; A21D; A23B,C,D,F,G,J, K,L; C12C,F,G,H,J; C13D,F,J,K
36 Petrochemistry, basic material chemistry	C09B,C,F,G,H,K; C10B,C,F,G,H,J, K,L,M; C11B,C,D

40 Process technology, special equipment	
41 Chemical process technology	B01B,D (excl.-046 - -053),F,J,L; B02C; B03; B04; B05B; B06; B07; B08; F25J; F26
42 Surface material technology, coatings	B05C,D; B32; C23; C25; C30
43 Material technology, metallurgy	C01; C03C; C04; C21; C22; B22
44 Processing of materials, textiles (*)	A41H; A43D; A46D; B28; B29; B31; C03B; C08J; C14; D01; D02; D03; D04B,C,G,H; D05; D06B,C,G,H,J,L,M,P,Q
45 Pulp and paper (*)	D21
46 Printing technology, packaging material	B25J; B41; B65B,C,D,F,G,H; B66; B67
47 Agricultural produce and food technology, machinery and equipment	A01B,C,D,F,G,J,K,L,M; A21B,C; A22; A23N,P; B02B; C12L; C13C,G,H
48 Environmental technology	A62D; B01D-046 - -053; B09; C02; F01N; F23G,J
50 Mechanical engineering, equipment	
51 Machine tools	B21; B23; B24; B26D,F; B27; B30
52 Engines, pumps, turbines	F01B,C,D,K,L,M,P; F02; F03; F04; F23R
53 Thermal engineering, processes and equipment	F22; F23B,C,D,H,K,L,M,N,Q; F24; F25B,C; F27; F28
54 Mechanical components	F15; F16; F17; G05G
55 Transport equipment	B60; B61; B62; B63B,C,H,J; B64B,C,D,F
56 Space technology, weapons technology	B63G; B64G; C06; F41; F42
60 Consumption goods and equipment	A24; A41B,C,D,F,G; A42; A43B, C; A44; A45; A46B; A47; A62B,C; A63; B25B,C,D,F,G,H; B26B; B42; B43; B44; B68; D04D; D06F, N; D07; F25D; G10B,C,D,F,G,H,K

70 Earth construction and hydraulic engineering, mining	E01; E02; E03; E04; E05; E06; E21
80 Software	(not IPC-class compatible)
81 Applications software	
82 Artificial intelligence	
83 Databases	
84 Data processing	
85 Security technology	
86 Data management systems	
87 Network software, network management	
88 Programming and programming languages	
90 'Problems'	
91 Ambiguous case	
92 Classification not applicable (service etc.)	
99 No information	

Sources:
10–70 Fraunhofer ISI / Jan 17, 1997 (* = own classification)
80 Vereinigung der Technologiezentren Österreichs:
http://www.tcs.co.at/vtoe/firmen/tcc/tcc.htm
IPC-classification: http://www.wipo.int/eng/clssfctn/ipc/ipc6en/index.htm

APPENDIX 3: Estimation results for the time period dummies

	Pooled Probit 1			Pooled Probit 2		
	Coef.	Robust Std. Err.	Partial effect	Coef.	Robust Std. Err.	Partial effect
Time periods (ref. POST1997)						
PRE1986	0.5070**	0.2030	0.1642**	0.6651***	0.2129	0.2465***
YEARS86-87	0.0651	0.2523	0.0238	0.1370	0.2531	0.0545
YEARS88-89	0.2361	0.2218	0.0830	0.4057	0.2550	0.1573*
YEARS90-91	-0.0235	0.2235	-0.0087	0.1484	0.2325	0.0590
YEAR1992	0.3408	0.2166	0.1163	0.6260***	0.2393	0.2339***
YEAR1993	0.2670	0.2181	0.0931	0.4594**	0.2184	0.1767**
YEAR1994	0.3250	0.2256	0.1114	0.4941**	0.2329	0.1891**
YEAR1995	0.0450	0.1992	0.0165	0.1966	0.2168	0.0779
YEAR1996	-0.0139	0.2039	-0.0052	0.0757	0.2079	0.0302
YEAR1997	0.0022	0.2036	0.0008	0.1651	0.2234	0.0655