[614] Working Paper

Lovely but dangerous: The impact of patent citations on patent duration

Per Botolf Maurseth

No. 614 March - 2001

Norwegian Institute of International Affairs
Norsk Utenrikspolitisk Institutt

Utgiver: NUPI

ISSN: 0800 - 0018

Besøksadresse: Grønlandsleiret 25
Addresse: Postboks 8159 Dep.
0033 Oslo
Internett: www.nupi.no
E-post: pub@nupi.no
Fax: [+ 47] 22 17 70 15
Tel: [+ 47] 22 05 65 00

Lovely but dangerous: The impact of patent citations on patent duration

Per Botolf Maurseth*
(The Norwegian Institute of International Affairs)

* Email: PerB.Maurseth@nupi.no. This paper has benefited from comments from Karl Ove Moene, Jan Fagerberg, Mads Greaker and Leo Andreas Grünfeld. I thank Tor Jakob Klette and Bart Verspagen for providing the data used in this paper.

[Abstract] What is the impact of patent citations on patent renewal behaviour? Patent citations are commonly used as an indicator of technology spillovers. For cited patents therefore, patent citations have a potentially ambiguous impact. On the one hand, patent citations may indicate a scientific breakthrough, a high value of the cited patent and therefore a long survival period. On the other hand, patent citations may indicate competing innovations that render the cited patent obsolete. By discriminating patents by technology field, it is demonstrated that patents that receive citations across technology fields survive longer than other patents. Patents that receive citations within the same technology field lapse earlier.

JEL Classification Numbers: 030, 034, C41
Keywords: Technology spillovers, Patent data, Patent Renewal Data, Patent Citations.

1. Introduction

It is widely accepted that spillover effects associated with knowledge production reduce the (private) incentives to invest in research and development, R&D.[1] The patent institution is the political answer to knowledge spillovers (in addition to R&D subsidies of various kinds). Patents exist in order to protect the property rights to innovations. Despite their apparent importance, little is known about the influence of technology spillovers on the value of innovations. It is recognised that the size distribution of private returns from technological innovation is skewed to the right (Scherer, 1998). While most innovations are of marginal economic value, some are extremely valuable. The same applies to the complete social value of innovations defined as the private value plus the value of pecuniary and purely external effects. According to recent literature on *general-purpose technology* (see e.g. Helpman, 1998) most innovations build on and refine existing technology. Yet, some innovations are path-breaking and open windows of opportunities for subsequent incremental innovations. Therefore, the private and social values of an innovation are likely to vary with the degree to which it relates to preceding innovations and the degree to which it stimulates subsequent innovations.

Empirical studies of the value of innovations have been based on a variety of approaches. For studies of the private value of innovations, indirect measures, like estimating firms' market value or profits as functions of innovations have been most frequent (see e.g. Griliches *et al.*, 1987 and Hall, 1993). These studies are based on econometric estimation using firm level data and either patents or research and development (R&D) as innovation indicator. Another approach is to study firms' patenting behaviour directly (see Pakes, 1986 and Pakes and Schankerman, 1986). In order to keep patents in force, holders of patents must pay an annual renewal fee. If patent-holders renew their patents based on an assessment of the value of the patent rights, data on patent renewal contain information on the distribution of the value of patents. This is the approach taken in this paper.

[1] See the textbooks by Barro and Sala-I-Martin (1995) and Aghion and Howitt (1998) for thorough discussions on the effects of knowledge externalities on economic growth. Benassy (1998) and Cohen and Levinthal (1989) provide arguments against the common view that incentives to investments in R&D are necessarily sub-optimal in decentralised market economies.

Studies of the total *social* value of innovation are also based on different methods. One approach is to estimate production functions including external R&D (i.e. R&D performed by other firms) (see e.g. Jaffe, 1986 or the survey by Griliches, 1992). A more recent approach is to use patent citations as a direct measure of R&D spillovers. Patent citations are added to patent documents when the intrinsic knowledge in the cited patent is relevant for the knowledge patented by the citing patent. In recent empirical work, the influence of patent citations on productivity in the *citing* sector (Verspagen, 1997) or on the value of *cited* patented innovations has been investigated (Hall *et al*., 2000).

In this paper, I use patent renewal data to draw inferences on the value of patents. While the existing literature has shed considerable light on the unconditional distribution of the value of patents, it has not to the same extent *explored* the factors behind the distribution by including explanatory variables in the analyses. This paper is devoted to the impact of patent citations on patent duration and therefore the value of patents. It is found that patent citations influence on the survival period of the cited patent, in the way predicted by theory: In general, patent citations correlate positively with survival time. Patent citations within technology classes have the opposite effect. Such citations may therefore indicate rival patents that creatively erode the value of the cited patent.

The next section provides a brief presentation of the patent institution and the information contained in patent documents. Also, the existing literature on patent citations, the value of patents and patent duration data are briefly reviewed. Section 3 presents a general model of patent renewal. Thereafter the database on patent renewals used in this paper and the empirical specification used are discussed. Section 5 provides empirical results. This section presents nonparametric, semiparametric and parametric results on patent survival and some estimates on the distribution of the value of patents. Section 6 concludes.

2. Patents, patent duration and patent citations

The patent institution

A patent is a document provided by legal authorities that gives the holder the exclusive rights to commercial exploitation of the described innovation for a certain period. The holder of a patent is therefore entitled to deny others to produce, trade or in other ways commercially exploit the innovation without explicit permission from the patentee. A prerequisite for patent protection is that the innovation has a potential commercial value. Purely scientific innovations are not patentable. It is also required that the patented innovation is new. In addition, the innovation has to be non-trivial and differ substantially from existing technology. All patent applications are public documents independently of whether the patent is granted or not.

Before a patent is granted, the patent office performs a search in existing patent documents in order to check that the idea to be patented is genuinely new. During this search period, patent citations (and citations to scientific literature) are added to the patent documents. Patent citations indicate relevant patented knowledge and they have the legal purpose to potentially limit the scope of patent protection. During the consideration, the patent examiners also assign the patent to one technology field according to the international patent classification system (IPC).

In most countries, patents are granted for a potential period of 20 years. The period of patent protection is potential because the granted patent has to be renewed annually. Annual renewal fees are positive, relatively low and progressive in the duration of the patent.

When a patent is granted in one country, it is valid in this country only. If an innovator wants his patent to be valid in more than one country, she has to apply for patent protection in all relevant countries. In Europe, the *European Patent Office* (EPO) has the authority to grant patents for all member countries or for a limited number of countries if the applicant wants so.[2]

Patent renewal, the value of patents and patent citations

Patent data are known for their deficiencies as a measure of innovation (see e.g. the survey by Griliches, 1990). In particular, simple patent counts neither take into account differences in the quality of innovations, the fact that many patents do not lead to commercial innovations nor the fact that the propensity to patent varies between sectors. Also, use of patent data does not take into account the fact that many patents are applied for for other reasons than appropriating the returns from innovations.[3] The benefit of patent data in research on technological change is the

[2] All EU member states are members of EPO in addition to Switzerland, Cyprus, Liechtenstein and Monaco. Norway is not a member of EPO.

[3] See Levin et al. (1987) for a study of the efficiency of patents as a means to appropriate the returns from innovation.

nature of the data. The very disaggregated level of information revealed in patent documents makes patent data a potentially important source of information.

In order to discriminate between patents of different economic impact, Pakes and Schankerman (1984) and Pakes (1986) analyse the survival of patents. The hypothesis is that patents that are renewed for longer periods have a higher economic value than patents that are renewed for short periods only. Later contributions are Pakes and Simpson (1989), Lanjouw (1998), Scankerman (1998) and Lanjouw *et al.* (1998). The typical finding in these studies is that the value distribution of patents is highly skewed with a median far below the mean. The majority of patents have a low value while a few are extremely valuable.

Patent citations have been used as an indicator of technology flows in a series of studies. This is because patent citations indicate a link of relevance between the cited and the citing patent. Jaffe and Trajtenberg (1998, p. 8) write that 'the appearance of a citation indicates that the cited patent is, in some sense, a technological antecedent of the citing patent'. From a survey study, Jaffe, Fogarty and Banks (1998) conclude that citations provide information about the generation of future technological impacts of a given invention.

Other contributions have explicitly interpreted patent citations as spillovers flowing from the cited to the citing patent. The flows of spillovers over geographical distance (Jaffe *et al.*, 1993, Sjøholm, 1996 and Maurseth and Verspagen, 1999), across country and language borders (Jaffe and Trajtenberg, 1998 and Maurseth and Verspagen, 1999), across sectors and technology fields and between different types of institutions (Jaffe and Trajtenberg, 1996) have been analysed. Jaffe and Trajtenberg analyse patent citations on a panel data set and study the time dimension in addition to the cross-country dimension. The studies suggest that patent citations are generated from a 'gravity like' process in which patent citations links are most intensive between agents being 'close' to each other. In the time dimension, citations links are most intense after a time lag of three to five years (Jaffe and Trajtenberg, 1996 and 1998).

Verspagen (1997) and Maurseth (2001) have investigated the impact of patent citations in the observational unit *citing* other units. These studies suggest that firms or regions that cite other firms or regions experience higher growth (in productivity or per capita income).

The above studies support the idea that patent citations indicate a positive spillover effect from the cited to the citing patents. How is the opposite effect? What is the impact of subsequent patent citations on a cited patent? Theory suggests that the effect can run in both directions: Patent citations received may indicate that the cited patent has an extra value above the average. *Ceteris paribus*, it seems likely that technologically pathbreaking and valuable patents are cited more often than other patents. On the other hand, patent citations might also indicate the well-known *creative destruction* effect in research rivalry. When a patent occurs that renders an existing patent obsolete, it seems likely that it will trigger a patent citation. Thus, patent citations might both indicate an extra value of a cited patent and intense competition towards the same patent. In empirical research, the second of these effects has not been taken into account. Trajtenberg (1990) analysed the social value of innovations on CT-scanners using hedonic price indexes. He concludes that particularly valuable innovation in this specific technology field received more patent citations than less valuable innovations. Patent citations can therefore be used as an indicator of the *social* value of innovations. Because spillovers are harmful for innovators it is not obvious that patent citations signal private value. Harhoff *et al*. (1999) find that the number of citations received is positively correlated with estimates of the private value of the cited patent in a sample of particularly valuable patents. Lanjouw and Schankerman (1999) construct an index for the quality of patents, and find that citations contribute positively to this index. Hall *et al*. (2000) investigate the market value of a sample of firms as function of both R&D, raw patent counts and patent counts weighted by patent citations. They find that patent citations add positively (but in limited amounts) to the market value of firms in addition to resources devoted to R&D and patent counts. Neither of these studies takes into account the possibility that patent citations may indicate the negative spillover effect mentioned above. To the best of my knowledge, there has been no attempt at discriminating between the creative destruction effect and the extraordinary economic value one assumes that patent citations signal in the literature. It is the aim of this paper to add some insights on these ambiguous effects of patent citations.

3. Patent renewal and the value of patents

The models sketched in this section are versions of the ones presented in Pakes and Schankerman (1986) and Pakes (1986). Consider an agent who holds a patent. In order to keep the patent in force the agent has to pay the renewal fee, C_t.[4] If the fee is paid, the agent earns the current return to patent protection, R_t. The return to patent protection measures the value of keeping the patent in force. Therefore, it is not a direct measure of the value of the patented innovation. Some innovations are not patented, among other reasons, because patenting involves disclosure of underlying ideas.[5] Other innovations could have a high value independently of patenting. Even so, it is reasonable to assume that the returns to patent protection are correlated with the returns to the patented innovations.

The agent is assumed to evaluate the expectation of the integral of discounted net returns, $E(R_{ô}-C_{ô})e^{-rô}dô$, where r is the discount rate, over the statutory limit to patent protection, T, and choose a life span for the patent, T*, in order to maximise the present value of expected net returns. The agent may stop paying the renewal fee. In this case the patent lapses forever and the return is equal to zero thereafter. To stop paying the renewal fee is therefore an absorptive state. Alternatively, the agent may continue to pay the renewal fee, in which case the patent holder earns the net rate of return (R_t-C_t) (which may be positive or negative) and also the option to renew the patent at later stages.

The case of certainty

To set ideas, assume that the agent knows the path of the rate of return, R_t, with certainty. The agent's decision problem is then to maximise the discounted value of net returns by choosing the age at which to stop paying the renewal fee. Therefore, the agent chooses a lifetime, T*, in order to solve the problem:

1) $$V = \max_{T^* \leq T} \int_0^{T^*} (R_t - C_t) e^{-rt} dt$$

[4] Here the renewal fees are assumed to be a function of the patent's age, t. The renewal fee schedules change from time to time, however, so an exact formulation would be to let the renewal fee depend on the patent's cohort as well. In the appendix, the Norwegian renewal fees are discussed and it is demonstrated that an age-specific, cohort independent fee schedule is an acceptable empirical approximation.

[5] See e.g. Levin et al. (1987)

In equation 1), V is the value of patent protection, given the optimal renewal behaviour. A simple renewal rule is to choose to stop paying the renewal fee when net return becomes negative, or if no such point of time exists, the maximum life span, T. A necessary condition for the implied life span to be optimal is that there exists a T* for which $(R_t-C_t)>0$ for t<T* and $(R_t-C_t)<0$ for t>T* in the neighbourhood of T*. For this life span to be optimal it is sufficient that (R_t-C_t) is non-increasing in t. If returns to patent protection decay by the annual rate ä, the vector **x** constitutes explanatory variables and β their coefficient vector, the renewal decision will be to renew patent i as long as:

$$2) \qquad R_{it} = R_{i0}e^{\mathbf{x}_i b} e^{-dt} \geq C_t \Rightarrow R_{io} \geq C_t e^{-\mathbf{x}_i b} e^{dt}$$

In equation 2) R_{i0} represents the initial return to patent protection for patent i. In this case, therefore, the costs schedule for patent protection, the explanatory variables and the survival length distribution are enough to draw inferences on the distribution of the value of patents. Let $g(R_{i0};\ \theta)$ and $G(R_{i0};\ \theta)$ be the density and cumulative distribution functions of initial returns of patents, where θ represents a parameter vector. Then the probability that patent i has dropped out by age t is given by:

$$3) \qquad Q_{it} = \int_0^{z_{it}} g(R_{i0};q)dR_{i0} = G(z_{it};q)$$

$$z_{it} = C_t e^{-\mathbf{x}_i b} e^{dt}$$

Given a distribution, g, of initial returns, equation 3) provides a basis for studies of patent duration. Q_t is a function of age, so it also serves as the cumulated probability function of patent duration, F(t). Its corresponding density function is f(t)=dF(t)/dt. The implied survival function is given by S(t)=1-F(t), and the hazard function by h(t)=f(t)/S(t). The survival function is the probability that a patent survives at least to age t. The hazard function indicates the rate at which patents lapse at age t, given that they have not lapsed before.

The case of uncertainty

It is unrealistic to assume that the sequence of net returns of patent protection is known to the patentee in advance. Before innovation, technological progress must be characterised by *genuine uncertainty*.[6] It is reasonable that the net returns from patented innovations are uncertain even after the patent is applied for. A patentee may be uncertain about future innovation that may render his own innovation obsolete, the dynamics of future prices in the relevant markets, the future costs of developing the innovation and about production costs for the final product. Thus, even if the patentee may feel certain about the time path of the renewal fees, C_t, the exact time path of R_t is hardly certain even when R_0 is known. It is reasonable that a patentee takes into account the uncertainty of future R_t in decisions on whether to renew a patent. Without imposing more assumptions on the problem it is problematic to draw clear conclusions. Uncertainty may go in either direction. At a point of time, events may occur that change the returns to patent protection by a small amount both in the positive and the negative direction. Also, new drastic innovations may occur that destroy the value of the patent.

Take the second type of event first. This is the kind of creative destructive innovation that drives the growth models of Aghion and Howitt (1992) and Klette and Griliches (2000). Innovations, and in particular patents, are known to come as events with a random time span between them. Often, the arrival of patents is modelled as a Poisson process in which patents occur at a known arrival rate , ë. Then the probability that no drastic innovation has occurred at age t is $e^{-ët}$. If so, the value of the patent at t is R_t. The probability that at least one drastic innovation has occurred is 1- $e^{-ët}$. In this case the value of the patent at t is 0. If the process of drastic innovations is Poisson, its effects can be modelled as a proportional addition to the decay rate of the value of patents.

Uncertainty is not only over arrival of drastic competing innovations. Uncertainty may also be present over market conditions. In this case, the expected change of the value of patents may be zero, but uncertainty might still have important effects on the decision to renew the patent. The holder of the patent might want to renew the patent even if the current rate of return is negative. Define the net rate of

---

[6] Research is – by definition – characterised as search for something unknown. Therefore there is no known probability function for all characteristics of what will be found. There might be a known

return, $P_t=R_t-C_t$ and consider the expected value of patent protection at any age, t, as a function of P_t and age, $V(P_t,t)$. The value of patent protection is then the expected discounted integral of net return over the entire possible lifetime for the patent, i.e. until the statutory limit, provided this number is positive:

$$4)\qquad V(P_t,t)=\max\left(0, E_t \int_{t=t}^{T} P_t e^{-rt}\,dt\right)$$

The value of patent protection at age t can be formulated as the rate of net return in the next instance plus the expected value of patent protection shortly thereafter if this sum is positive:

$$5)\qquad V(P_t,t)=\max\left(0, P_t dt + e^{-rdt} E\left[V(P_t+dP, t+dt)\middle| P_t\right]\right)$$

Equation 5) is the Bellman equation of the optimal stopping problem that faces the holder of the patent. The stopping problem implies, in opposition to the renewal rule outlined above, that patents may be renewed even if their current net rate of return is negative. The holder of the patent may want to renew the patent because rates of return may turn positive on later stages. The problem described in equation 5) has been analysed on a general basis in several contributions (see e.g. Pindyck, 1991, Dixt, 1992 and 1993 and Dixit and Pindyck, 1994). For some stochastic processes of P, the problem results in a threshold rule for renewal, P(t)*. Generally, this rule will be to renew the patent as long as P>P(t)* and to let the patent lapse whenever P<P(t)*. P*<0, so the threshold net return of patents (above which the patent will be renewed) is negative. P* will also increase in time towards zero at t=T. This follows from the fact that the option to renew the patent in later periods disappears at the end of the statutory period. The existence of the threshold value is ensured when there is positive persistence of uncertainty so that the cumulative probability distribution of future net rates of return, P', H(P'|P) shifts to the right when P increases.[7] When this

probability function of the *value* of the results of research, but as described in section 2, this probability function may not be very restrictive.

[7] See Dixit and Pindyck (1994) and Dixit (1993).

is the case, the probability that returns in future states are larger than any given number is larger the higher are current returns.

When future returns are uncertain and obey the assumptions described, there will exist a probability distribution function for lapse dates and therefore a density function, a survival function and a hazard rate. The probability function (and the implicit density, survival and hazard functions) will be conditional not only on initial returns but also of the stochastic process generating subsequent returns (Pakes, 1986).

The case of patent renewal under uncertainty is analysed by Pakes (1986) and Lanjouw (1998). Pakes and Simpson (1989) introduce 'stochastical dominance' as an approach to studies of patent duration. A class of patents, i, is said to stochastically dominating another, j, when the empirical survival function S*(t) of group i is larger than for group j, $S^*_i(t)>S^*_j(t)$. When class i dominates j, the proportion of patents in class i that have value greater than any number is larger than the corresponding proportion of patents of class j. This approach will be used in the non-parametric specification described in the next section. For the parametric specifications, I concentrate on the myopic case corresponding to the case of certainty.

4. Empirical specification and the data used

Censoring and data characteristics

The renewal data are from Statistics Norway, collected from the Norwegian Patent Office. They cover all patent applications in Norway during the period 1980-94 (in total more than 23 000). These data include both patents applied for by Norwegians and patents applied for in Norway by foreigners. The data include the patents' publication numbers, the application dates, the dates the patents were granted and the dates they lapsed (i.e. the dates at which the holder stopped paying the renewal fee). The time span between date of application and date of lapse is used as the observed span for uncensored observations. For censored observations, the time span between date of application and the latest observed lapse date is used. The latest observed lapse date is in 1992, so this shortens the period covered.

The data include the technology field of the patent according to its classification by International Patent Class (IPC). IPC is a hierarchical eight-digit classification system for technology fields of patents. IPC classes are defined according to technology, not economic sectors. In principle, therefore, a patent in pharmaceuticals can be classified in the same IPC code as a patent in chemicals. Verspagen *et al*. (1992) provide a concordance table between IPC and ISIC. Based on this concordance table, each patent was assigned one industrial sector.[8] The industrial sectors were further classified into four different groups according to R&D *intensity* (R&D as share of value added) and one common group for which R&D intensity could not be constructed.[9] The four groups are included as dummy variables in the estimated equations.

The renewal fee scheme was collected from the Norwegian Patent Office.[10] The renewal fees are increasing in age and the scheme changes from time to time. These changes are minor in terms of real value, so I imposed a common renewal fee for all cohorts based on the computed average of renewal fees in real 1980 Norwegian Kroner (NOK), as described in the appendix.[11]

The Norwegian patent data do not contain patent citations. The data on patent citations are from the European Patent Office (EPO). The database covers all European patents applied for in the period from 1979 to -96 and patent citations between them. The European patent database include the patents' publication numbers, information on IPC, the application dates but not lapse dates. Norway is not a member of EPO. Therefore, only Norwegian patents also applied for in EPO member countries are covered by both databases. These are few observations. In total 737 patents were identifiable in both databases. Of these only 598 could be included in the analysis since some were not granted before the latest observed lapse date.

There are two sorts of censoring in the data. The first type is for the survival data. The data cover Norwegian patents applied for in the period from 1980 to -94. Therefore, even the oldest patents do not reach the statutory limit to patent protection (which is 20 years in Norway). Consequently, all patents in the data that did not lapse are censored. The latest observed date of lapse in the data is in 1992. All patents that

---

[8] The industrial classification system used is ISIC rev. 2.The concordance scheme is not perfect. Some IPCs are assigned several ISIC sectors indicated by particular percentage shares. The patents were assigned to the sector that received the highest percentage. When the percentage was equal to 50, the first sector was chosen.

[9] R&D data are from the OECD ANBERD database. Data on value added are taken from the OECD STAN database. R&D intensity was averaged over the period studied and the 22 sectors were aggregated into the four sectors according to R&D as share of value added. Two sectors could not be assigned R&D intensities and these two are aggregated into a common subgroup. These are *building and construction* and *utilities* (sector 4000 and 5000, ISIC rev. 2) for which R&D is not reported in the ANBERD database and *other transport* (sector 3840 except –41, -43 and –45) for which no production in Norway is reported in the STAN database.

[10] Patentstyret (1979, 1982, 1986, 1993).

did not lapse before 1992 are consequently treated as censored. Without censoring, estimation of the probability density function of drop outs from patent protection would be to maximise the likelihood function

$$L(q) = \prod_{i=1}^{n} f(t_i, q)$$

Since the data are characterised by censoring, this has to be taken into account. Patent i is observed over the span $[c_i, \min(c_i+T^*_i, c_i+T')]$, where c_i is the application date and T' the latest observed lapse date. Let d_i be a dichotomous variable where di=0 denotes censoring and $d_i=1$ failure. The contribution to the likelihood for patent i is therefore given by (in which the parameter vector is suppressed):

10) $$L_i = f\left(\left(c_i + T_i^*\right)\right)^{d_i} S\left(\left(c_i + T_i'\right)\right)^{(1-d_i)}$$

The data on patent citations are also censored. The patent citations data cover the period from 1979 to 1996. In principle, a patent can be cited any time after its application, also after it lapses. Therefore, the number of citations any patent receives will be a function of whatever rule the patent citations possess (and therefore their relation to the value of patents) and in addition the time span from application to 1996. A patent that occurs late in the period observed (1980 to 1992) will on average receive fewer citations than a patent that occurs early in the period, if their value (and other characteristics) are equal. Jaffe *et al*. (1998) argue that the time profile of patent citations is a double exponential. Without taking the time profile of the present citations data into account, I have chosen to follow the approach in Trajtenberg (1990) and weight the number of patent citations received by the linear time span from application to 1996. Therefore, the explanatory variable included is based on the number of citations received per day since application of the patent.[12] In order to better interpret the estimates, this number is multiplied by 1825 (the number of days in a five years period). The resulting estimates therefore indicate the effect for

[11] One 1980 USD corresponds to 4.95 1980 NOK.

[12] Jaffe et al. argue that patent citations occur most frequently after a period of 2-4 years. For late-occurring patents, therefore, the choice of the linear weighting scheme described above will underestimate the impact of patent citations on patent duration.

patents that receive one patent citation every fifth year (which is somewhat below the average for cited patents).

For citations *within* the same technology field, the underlying hypothesis is that these potentially trigger lapse. For these patent citations, therefore, I weight the number of citations by the linear time span from application to the date of lapse (or censoring). The explanatory variable for this kind of patent citations is specified as the number of citations per day since application to the latest observed point of time before lapse or censoring multiplied by 1825. For patents receiving citations within the same technology field one such citation every fifth year is somewhat above the average.[13]

The implicit hypothesis for the construction of the explanatory variables is that *all* (both within and between technology classes) patent citations a patent receives potentially indicate a high value. Potentially dangerous citations are assumed to be within technology class citations. These are allowed to influence only during the patents' lifetime.

The 598 patents are not representative for the total sample of patents. Figure 1 shows the Kaplan-Meier survival estimate of the foreign patents (NOP=0), the Norwegian patents that were not applied for in EPO (NOP=1) and the Norwegian patents also identified in the EPO database (NOP=2). Analysis time is days from application of the patents.

<u>Insert figure 1 here</u>

The figure reveals three notable characteristics of the data. The first is that for all the three categories of patents, there is great variance of survival time and therefore of the implicit value of the patents. Less than 50 per cent of the patents in any category survived more than 4000 days (about 11 years). Secondly, there is a notable difference in renewal behaviour between Norwegian patents and foreign patents in

[13] An alternative specification would be to let inter-technology class citations enter the decay rate. The effects of such citations will be characterised by lags (both forward and backward) and intra-class citations in some technology classes will have different effects from citations within other classes. Also, even if one interpretation is that patent citations within a certain technology class trigger sudden death, another interpretation is that patents that receive such citations are subject to higher competition and therefore have lower value than patents that do not, also at times when no citations occur. Therefore, I decided not to use time-varying decay rates, but rather to let intra-class citations impact on the conditional mean survival time.

the data. At all ages, the foreign patents dominate the domestic ones. Thirdly, the Norwegian patents also applied for in EPO dominate the other domestic patents.[14]

Figure 2 shows the frequency of patent citations. The figure shows that patent citations have a very skew distribution. For the 598 patents used her, more than 85 per cent are never cited. Only one patent receives the maximum number of eight citations.[15]

<u>Figure 2</u>

It is interesting to note that patent duration does not seem to vary with R&D intensity. In figure 3, Kaplan-Meier estimates of survival are graphed according to the five subgroups of R&D intensity. The out-lying survival curve is for the subgroup for which R&D intensity could not be constructed. For the four other subgroups, no significant differences are visible.

<u>Figure 3</u>

Non-parametric testing and graphical approaches.

A common approach in analysing survival data is to use the strictly empirical and non-parametric *Kaplan-Meier* estimator. It does not depend on any restriction imposed by the model and it takes into account censored observations (observations that end because the observational period ends and not because the patent lapses). Let h_k be the number of spells completed (i.e. the number of patents that lapsed) at age k. Let m_k be the number of censored observations at age k (i.e. the number of patents that are not observed for higher ages because the dataset is censored). Let n_k be the number of patents that neither lapsed nor were censored before age k. These are therefore patents that lapse or are censored at later stages: $n_k=\Sigma_{l\;k}(m_l+h_l)$. Therefore, n_k is the number of patents that are 'at risk' at age k. The corresponding

[14] When survival was counted from date of grant instead of application, the difference between Norwegian and foreign patent application disappeared. The difference between the two Norwegian subgroups remained. This indicates that it takes longer for foreigners to get a patent application granted.

[15] The finding that patent citations are skew-distributed is very typical. See Jaffe *et al.* (1993), Trajtenberg (1990) or Maurseth and Verspagen (1999).

Kaplan-Meier estimator, or product limit estimator for the survivor function is then given by:

6) $$\hat{S}^{*}(t_k) = \prod_{l=1}^{k} \frac{(n_l - h_l)}{n_l}$$

This estimator implies that the estimated probability of a spell being completed at age t_k is set equal to the observed relative frequency of completion at this age. The Kaplan-Meier estimator can be used to compare survival performance among different categories or groups of the data. Differences among groups of the data can be tested using the so-called *log-rank test* developed by Kalbfleisch and Prentice (see Kiefer, 1988 or StataCorp, 1999). Inferences on differences in value between groups of patents will therefore be based on testing the significance of the difference between their survival functions. Use of the Kaplan-Meier estimator can not take into account continuous characteristics of the data, however.

Semi-parametric specification

In order to take into account continuous explanatory variables, some assumptions have to be imposed on the model to be tested. For the purpose of this paper, the empirical specification follows Pakes and Schankerman (1984 and 1986) and Schankerman (1998) in which uncertainty is disregarded. The results should therefore be interpreted as a reduced form formulation for the effects of the covariates and the estimates of initial returns, the decay rate and the other parameters as crude approximations. For survival data, a compromise between strict parametric models and non-parametric testing is the so-called Cox model. The Cox model presumes a constant *baseline hazard* that does not vary between individuals. Explanatory variables are taken into account as influencing the hazard rate directly. The hazard function to be estimated is therefore:

7) $$\mathrm{h}(\mathrm{t}|\mathrm{x}_\mathrm{i}) = h_0(t)\exp(\mathbf{x}_i b)$$

$\mathbf{x}_i$ are explanatory variables included and β is the coefficient vector.[16] The relative hazard between two observations is assumed to be independent of the baseline hazard and estimation can be performed without any other functional restrictions. It is a restriction though, that the baseline hazard is assumed to be equal for all observations and proportional over time. For some probability functions of initial returns, like the log-normal to be formulated below, the proportional hazard assumption is not valid.

Parametric approaches

In the literature, three alternative specifications of f(·) have been explored: the Pareto, the exponential-Weibull and the log-normal distributions. Pakes and Schankerman (1984) find that the Pareto distribution fits an international dataset better than the other two, while Pakes and Schankerman (1986) argue that the log-normal distribution fits a newer, but essentially similar dataset somewhat better. Schankerman (1998) obtains a similar result based on French data. In the Pareto case, the density function of survival for patents will be exponential, in the exponential-Weibull case, the density function becomes a Gompertz-like function, while in the log-normal case, the density function is normal. I experimented with all the three distributions and the log-normal out-performed the two others in two different ways: Firstly, for neither the Pareto nor the exponential-Weibull distributions, the estimated parameters were economically sensible (they implied that the renewal rate increased with the renewal fees).[17] Secondly, the fits of the model, both in terms of predictions and in terms of the resulting likelihood, were better for the log-normal distribution than for the two other distributions. The deduction of the density and survival functions for the Pareto distribution and the exponential-Weibull distribution are presented in the appendix.

If initial returns distribute log-normally, the log of initial returns, $r_0=\ln(R_0)$ distributes normally, $r_0 \sim N(\mu, \sigma)$. In this case, an agent will renew a patent of age t as long as log of initial returns is given by:

[16] In equation 7), costs are excluded. The reason is that costs influence on the hazard function in a restricted manner which has to be estimated.

[17] These results are similar to those obtained by Schankerman (1998).

8) $r_{0i} \geq \ln C_t + dt - \mathbf{x}_i b$,

or, equivalently the standard normally distributed expression :

$$\frac{r_{0i} - m}{s} \geq \frac{-m + \ln C_t + dt - \mathbf{x}_i b}{s}$$

The age of the patent, t, enters in two of the terms in equation 8), in $\ln C_t$ and in ät. An approximation is to treat $\ln C_t$ as individual specific variables and solve directly for t from ät. Therefore, survival time will be given by:

9) $t_i = \frac{1}{d}\left(m - \ln C_t + \mathbf{x}_i b\right)$

Therefore, t distributes approximately normally. Because of censoring and non-linearity, it is difficult to estimate equation 9) directly. In order to obtain an estimate for ä, I follow Schankerman and Pakes (1986) and Schankerman (1998). Their procedure differed from the present paper since they did not include explanatory variables at the micro-level. In stead, their approach was to estimate the survival function based on the observed survival rates. To investigate the value distribution of patents, the point of departure is the empirical Kaplan-Meier estimates of the survival function, S*(t), i.e. the fraction of patents being renewed at least to age t. The proportion of patents that survive until age t is given by:

10) $S^*(t) = 1 - \Phi(z_t)$

I impose a common decay rate on all patents and I therefore suppress the individual explanatory variables in the following. In this case, the model can be written:

11) $$\begin{aligned} y_t &= 1 - \Phi^{-1}\left(S^*(t)\right) \\ &= \frac{-m + \ln C_t + dt}{s} + u_t \end{aligned}$$

Above, Ö denotes the cumulative normal distribution. The disturbance term u_t is a composite error term capturing errors in the renewal rule, w, and a binomial sampling error, v_t, with variance S*(t)(1-S*(t)/N, where N is the number of patents in the sample. The weights are estimated following Amemiya (1981) as described in the appendix, and equation 11) is estimated by weighted non-linear least squares.

With the estimate of ä, ä^ at hand, the density and survival functions of survival time can be expressed as in equation 12) and 13):

$$12)\qquad f(t_i) = \frac{\hat{d}}{s}\, j\left[\frac{t_i - \left(1/\hat{d}\right)\left(m - \ln C_t + \mathbf{x}_i b\right)}{s}\right]$$

$$13)\qquad S(t_i) = 1 - \Phi\left[\frac{t_i - \left(1/\hat{d}\right)\left(m - \ln C_t + \mathbf{x}_i b\right)}{s}\right]$$

In equations 12) and 13) notation is as above and ö denotes the standard normal density function.

5. Empirical results

Non-parametric results and testing

Graphs 4, 5 and 6 below present Kaplan-Meier graphs of survival of patents conditional on categorical indicators of patent citations. Graph 4 is for cited patents (CITED=1) and not cited patents (CITED=0). The graph seems to support the idea that cited patents are renewed for longer periods than patents that are not cited. The survival curve for cited patents is almost everywhere above the curve for not cited patents. This effect seems to be small, however. This result mimics those of Hall *et al.* (2000) who find clear, but only marginal evidence that patent citations increase the market value of the firm to which the patent belongs.

Graphs 5 and 6 present corresponding results on citations within and between technology classes. In order to discriminate between patent citations between and within technology classes, an explorative method was used. The IPC class is an hierarchical eight-digit code. Therefore, cut-offs between within and between IPC classes could in principle be constructed for differences in number up to the eighth digit. It would also be possible to make use of the industrial classification and analyse patent citations within and between industrial sectors. For the purpose of this paper, within-technology citations were defined as citations between patents that differed in the three last digits only. All other citations were defined as between-technology citations. The cut-off at the three last digits was chosen because it maximised the difference in the observed survival behaviour.[18]

Graph 5 presents survival estimates for the group of patents that receive inter-technology class citations versus other patents. Graph 6 presents similar evidence for patents that only receive intra-technology field citations.

<u>Insert figures 4, 5 and 6 here</u>

Figures 5 and 6 illustrate the main findings in this paper: The effect of patent citations on patent duration is consistent with the dual role of technology spillovers described in section 2. Patents that are cited across technology fields are renewed for longer periods than patents that are not cited (INTERC=1 versus INTERC=0). Patent citations in general seem to indicate that the cited patent is of particular technological value and such citations therefore indirectly imply a higher economic value of the cited patent. Looking at inter-technology class citations instead of all citations markedly strengthens the evidence in favour of this interpretation.

Patent citations within technology classes have the opposite effect. Patents that receive citations within the same technological class only are renewed for shorter periods than other patents (INTRAP=1 versus INTRAP=0). This finding supports the hypothesis that intra-technology field citations indicate rival innovations.[19]

<u>Table 1</u>

[18] Probably the optimal cut-off to discriminate between within- and between-technology class citations will vary between patent classes. In future research I intend to work further on this question and construct alternative cut-off criteria.

[19] It has been objected to me that the survival curves differ only for some ages. It should be underlined that the survival curves for subgroups and for the oldest ages are based on very few observations.

Neither the log-rank nor the so-called Wilcoxon tests reject equality between the survivor function graphed in figures 4 through 6. Test results are presented in table 1. It should be underlined that these results are obtained on a small sample and that the categorisation of the data suppresses details. The category of patents that receives inter-technology class citations includes patents that receive only such citations and patents that receive both types of citations. Furthermore, the patents receive different amounts of the two categories of citations.

Semi-parametric Cox model

To take into account the continuous characteristics of the explanatory variables, it is necessary to rely on (semi-) parametric methods. The simplest way to do this is to use the semi-parametric Cox model.

In the regressions, the main explanatory variables included are INTRA and CITES. INTRA denotes the number of intra-technology class patent citations per fifth year during the patents' lifetime as discussed above. CITES denotes the total number of citations a patent receives every fifth year during the time span from date of application to 1996, as discussed above. In addition to these variables, dummy variables for the years of application (up to 1988) and for the four categories of R&D intensity were included in the regressions.[20]

Given the limited number of citations in the sample, it is possible to explore another characteristic of the data. Some patent citations are *self-citations* in the sense that a patent receives citations by patents applied for either by the *same inventor* who developed the original patent or by the *same firm* that applied for the original patent. By searching through the entire sample of patent citations, 14 citations were identified as self-citations. Technology spillovers are external effects from one innovation to other innovators. As such, self-citations represent noise to the data. It is not clear, however, that the effect of citations on the value of patented innovation should differ according to whether they are self-citations or not. The *creative destruction effect* of intra-technology class patents could be present even if the innovator of a patent chooses to introduce a new vintage of a product. Similarly, the

[20] It was also experimented with inclusion of other explanatory variables, like the number of patents applied for within the specific sector or technology class during the whole period or the same year in order to investigate possible effects of competition. Neither of these variables resulted in significant coefficients. I also experimented with using sector-specific dummy variables, but no notable results were obtained.

extra value of a patent signalled by citations in general may be present even if it is the original inventor who adds citations. To allow for differences in the nature between self-citations and external citations, I report results both when self-citations were included in the common citations variables and when they were treated separately.

Table 2.

The results from the Cox regressions are reported in table 2. The coefficients are reported with exponentiated coefficients. The reported estimates can therefore be interpreted as the ratio of the hazards between patents that differ one unit in the corresponding covariate. For instance, the coefficient of INTRA is 6.59. Therefore, a patent with one patent citation within the same technology class every fifth year faces a hazard that is 6.59 times as high as a patent that receives no patent citations. This implies a very high hazard rate in the case of intra-technology class citations. The result is the opposite for citations in general. The general effect of patent citations, when the ‘dangerous’ within-technology class citations are accounted for, is very productive. A patent that receives one citation every fifth year has a hazard rate that is only 16 per cent of that of patents that receive no citations. When self-citations are taken into account, the effect of external patent citations is even stronger. Self-citations do not significantly influence on the estimated hazard rates, but the effect of the residual citations, external citations, is larger when self-citations are excluded from the variables.

It is interesting to note that the hazard rate does not seem to be increasing with R&D intensity in the sector to which the patents belong, although a systematic analysis of this aspect is beyond the scope of this paper. The interested reader should confer Schankerman (1998). The results do not reveal any clear time trend for survival of patents. As compared to 1980 (the base year) all coefficients are negative, but there is no trend in their changes from year to year.

The value distribution of patents and parametric estimation

Table 3 reports results from weighted non-linear estimation of equation 11. The dependent variable is based on the Kaplan-Meier survival estimates, to take into account that the data are censored. Neither cohort-specific nor sector-specific differences are included in the regressions. This would have reduced the number of observations dramatically.

Table 3

The results imply a very skew value distribution of patents. The time unit is days, so the estimated parameter μ (5.45) is the log of mean daily rates of return in 1980 NOK. The coefficient of δ is the estimate of how fast patent value decays. The implied yearly rate is 45 per cent. As compared to the result in Pakes and Schankerman (1984 and 1986) and in particular in Schankerman (1998) this is very high. Their estimates vary between 2 (one sector in Schankerman's 1998 article) and 25 per cent. Two differences between the present study and their study should be noted. Firstly, Pakes and Schankerman analysed samples covering the period from 1950 to 1980. The period studied here is from 1980 to 1992. Secondly, Pakes and Schankerman excluded the first five years from their study. In this paper, the whole period from application onwards is included. The estimated standard deviation is very similar to those obtained by Pakes and Schankerman (1986).

Table 4

Table 4 reports the implicit distribution of the initial value of patents by quartile. The table is constructed by drawing 50 000 observations from a pseudo-random normal distribution with the estimated initial mean rates of return and the standard deviation imposed. The resulting numbers are thereafter exponentiated in order to obtain the log-normal distribution. Then the resulting numbers are converted to yearly rates of return. The table confirms the typical finding from studies of the distribution of the value of patents. The mean exceeds the median by a large amount. Less than ten per cent of the patents have initial yearly rates of return in excess of 1 million 1980 NOK.

In order to investigate the impact of patent citations on both renewal behaviour and the value of initial returns, a log-likelihood function based on the density and survival functions presented in equations 11 and 12 above was constructed. The inverse of the estimate of δ was imposed on the log of the renewal fee schedule and constrained estimation was conducted. The results, both with and without self-citations accounted for, are reported in table 5.

Table 5.

The table confirms the very strong impact of patent citations reported in table 2. In contrast to that table, the coefficients in table 5 can be interpreted directly as the impact of a unit's change in the explanatory variable on the expected survival time for patents. One external patent citation every fifth year has an estimated impact of increasing the survival period by more than three years (1336 days). One intra-technology class patent citation from another agent each fifth year decreases the survival period by more than three years (1360 days). The effects of patent citations can be simulated and graphed against time. Figure 7 shows the implied survival functions for patents that do not receive patent citations, for patents that receive one citation every fifth year and for patents that receive one intra-technology class citation every fifth year. Figure 7 also graphs the estimated Kaplan-Meier survival function from the observed data.

Figure 7

There are three conclusions to be read out of the figure. Firstly, the model seems to fit the data reasonably well. There is a lack of fit for the oldest patents, though this is partly due to the omission of dummy variables for this specific regression.[21] Secondly, the impact of patent citations as described above is very clear. Patents that receive one intra-technology class citation every fifth year have markedly lower survival rates than other patents. Patents that receive the same amount of the general 'productive' citations have higher survival rates. Thirdly, the effect of the discretely changing renewal fee scheme is evident. The simulated survival curves have breaks in them, corresponding to the increase in renewal fees every third year. The

[21] The figure is constructed by rerunning the estimation without using dummy variables. On the basis of the obtained estimates, the respective survival functions are simulated. The reason for excluding the dummies is that inclusion imposes an arbitrary basis patent group for which the mean is reported. In table 5, the basis group is patents applied for in 1980 belonging to the sectors for which no R&D intensity could be constructed.

simulated curves do not capture the fact that renewal of patents occurs yearly, as visible in the stepwisely observed survival curves, however.

What are the conclusions from these estimates for the value of patents? The way the regression model is formulated, patent citations are assumed to influence on the conditional mean of the patents while the standard deviation and the decay rates are assumed to be unaffected and common for all patents. Therefore, the influence of patent citations on value is given as an addition of the estimated influence on days multiplied by the estimated decay rate to the parameters in the log-normal distribution. Patents in the basis group (i.e. patents applied for in 1980 and belonging to the sectors for which R&D intensities could not be constructed) had initial mean daily rates of return equal to NOK 4 357. Patents belonging to the same subgroup with one 'productive' patent citation every fifth year had expected initial daily rates of return equal to NOK 21 649. For a patent belonging to the same group receiving one 'dangerous' patent citation every fifth year during its lifetime, expected initial daily rates of return are reduced to NOK 852.[22]

The above calculations support the hypothesis put forth in this paper: Patent citations in general do signal higher value of the cited patent. When the citing patent is within the same technological 'neighbourhood' as the cited patent, the effect is the opposite. In this case, patent citations seem to erode the private economic value of the cited patent.

6. Conclusion and future research

It is well known that the value of innovation has a very skew distribution. This fact implies that determinants of the value of innovation are important research issues. In particular, the determinants of technological interaction are important. In previous research, patent citations have been used as an indicator of above-normal value of patents. The idea is that patents that receive many patent citations constitute major scientific breakthroughs that are useful for later research and also indicate economic extra value to the patentee. In this paper, this has been questioned. Generally, the evidence supports the above hypothesis. This interpretation is strengthened when the particular nature of citations within technological borders is accounted for. Such

[22] For the log-normal distribution, the median and the mean are given by $\exp(\mu)$ and $\exp(\mu+1/2\sigma^2)$. The calculations above are based on the estimated coefficients, multiplied by the estimated daily decay rate of 0.0012.

citations indicate shorter survival of the cited patents. These findings are in accordance with the interpretation of patent citations as knowledge spillovers. Knowledge spillovers *per se* are external effects from innovations. As such they reduce the (private) value of innovations. It is not arbitrary what patents which produce knowledge spillovers, however. Spillover-producing patents are of higher technological and economical value than other patents. When knowledge spills over to rival innovators, they signal competition that drastically reduces the (private) value of patents.

Research on patent citations is a new field. The present study will be extended in future research. Firstly, the impact of patent citations should be analysed on a more disaggregated level. The level of similarity in technological classes that determines whether a patent citation should be interpreted as a rival patent or an indicator of extra value should be analysed at industry level. Secondly, a full-fledged analysis of patent survival could introduce time-varying variables indicating not only the extent of, but also the time path of patent citations occurring. Thirdly, by analysing the time span of drastic innovations, the possibility for analysing the length of product cycles occurs. In order to take such advantages of patent citations, more detailed data is required.

References

Aghion, P. and P. Howitt (1998) *Endogenous Growth Theory*, MIT Press, Cambridge, Massachusetts.

Barro, R. J, and X. Sala-I-Martin (1995) *Economic Growth*, McGrawhill, New York.

Benassy, J. P. (1998) "Is there always too Little Research in Endogenous Growth with Expanding Product Variety?" *European Economic Review*, 42(1), pp. 61-69.

Cohen, W. M. and D. A. Levinthal (1989) "Innovation and Learning: The Two Faces of R&D", *Economic Journal* 99(397), pp. 569-96.

Dixit, A. K. (1992) "Investment and Hysteresis", *Journal of Economic Perspectives* 6, pp. 107-132.

Dixit, A. K. (1993) *The Art of Smooth Pasting*, Harwood Academic Publishers, Chur.

Dixit, A. K. and R. S. Pindyck (1994) *Investment Under Uncertainty,* Princeton University Press, Princeton.

Griliches, Z. (1992) “The Search For R&D Spillovers”, *The Scandinavian Journal of Economics* 94 (supplement), pp. 29-47.

Griliches, Z., B. H. Hall and A. Pakes (1987) “The Value of Patents as Indicators of Inventive Activity” in Dasgupta, P. and P. Stoneman (eds.) *Economic Policy and Technological Performance,* Cambridge University Press, Cambridge, pp. 97-124.

Hall, B. H. (1993) “The Stock Market Valution of R&D Investment during the 1980’s” *American Economic Review* 83, pp. 259 -264.

Hall, B. H., A. B. Jaffe and M. Trajtenberg (2000) *Market Value and Patent Citations: A First Look*, NBER Working Paper 7741, Cambridge, Massachusetts.

Harhoff, D., F. Narin, F. M. Scherer and K. Vopel (1998) “Citation Frequency and the Value of Patented Inventions” *Review of Economics and Statistics* 81 (3), pp. 511-515.

Helpman, E. (ed.) (1998) *General Purpose Technologies and Economic Growth* MIT Press, Cambridge, Massachusetts.

Jaffe, A. (1986) “Technological Opportunity and Spillovers of R&D: Evidence from Firms’ Patents, Profits and Market Value” *American Economic Review* 76 (5), pp. 984-1001.

Jaffe, A. B., M. S. Fogarty and B. A. Banks (1998) “Evidence from Patents and Patent Citations on the Impact of Nasa and other Federal Labs on Commercial Innovations” *Journal of Industrial Economics* XLVI (2), pp. 183-205.

Jaffe, A. B. and M. Trajtenberg (1996) *Flows of Knowledge from Universities and Federal Labs: Modelling the Flow of Patent Citations over Time and across Institutional and Geographical Boundaries*, NBER Working Paper 5712, Cambridge, Massachusetts.

Jaffe, A. B. and M. Trajtenberg (1998) *International Knowledge Flows: Evidence from Patent Citations*, NBER Working Paper 6507, Cambridge, Massachusetts.

Jaffe, A.B., M. Trajtenberg and R. Henderson (1993) "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations" *Quarterly Journal of Economics* 108, pp. 577-598.

Kiefer, N. M. (1988) "Economic Duration Data and Hazard Functions" *Journal of Economic Literature* XXVI, pp. 646-679.

Lanjouw, J. O. (1998) "Patent Protection in the Shadow of Infringement: Simulation Estimations of Patent Value", *Review of Economic Studies* 65, pp. 671-710.

Lanjouw, J. O. and M. Schankerman (1999) *The Quality of Ideas: Measuring Innovation with Multiple Indicators,* NBER Working Paper 7345, Cambridge, Massachusetts.

Lanjouw, J. O., A. Pakes and J. Putnam (1996) *How to Count Patents and Value Intellectual Property: Uses of Patent Renewal and Application Data,* NBER Working Paper 5741, Cambridge, Massachusetts.

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), pp. 783-831.

Maurseth, P. B. (2001) *Convergence, Geography and Technology* (forthcoming).

Maurseth, P. B. and B. Verspagen (1999) "Europe: One or Several Systems of Innovation? An Analysis Based on Patent Citations" in Fagerberg, J., P. Guerrieri and B. Verspagen (eds.) *The Economic Challenge for Europe – Adapting to Innovation Based Growth*, Edward Elgar, Cheltenham, pp. 149-174.

OECD (1996) *DSTI (STAN Industrial Database)* Pars.

OECD (1997) *DSTI (ANBERD)* Paris.

Pakes, A. (1986) "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks", *Econometrica* 54, (4), pp. 755-784.

Pakes. A. and M. Schankerman (1984) "The Rate of Obsolescence of Patents, Research Gestation Lags and the Private Return to Research Resources" in Griliches Z. (ed.) *R&D, Patents and Productivity*, NBER, University of Chicago Press, Chicago, pp. 73-88.

Pakes, A. and M. Simpson (1989) "Patent Renewal Data", *Brookings Papers of Economic Activity,* Microeconomics, pp. 331-410.

Patentsyret (1979, 1983, 1986, 1993) *Forskrift til patentloven* (Regulations of patent law), Oslo.

Pindyck, R. S. (1991) "Irreversibility, Uncertainty and Investment", *Journal of Economic Literature* 29, pp. 1110-1152.

Schankerman, M. (1998) "How Valuable Is Patent Protection? Estimates by Technology Field", *RAND Journal of Economics* 29(1), pp. 77-107.

Schankerman, M. and A. Pakes (1986) "Estimates of the Value of Patent Rights in the European Countries during the Post-1950 Period", *Economic Journal* 96, pp. 1052-1076.

Scherer, F. M. (1998) "The Size Distribution of Profits from Innovation", *Annales d'Économie et de Statistique* 49/50, pp. 495-516.

Sjöholm, F. (1996) “International Transfer of Knowledge: The Role of International Trade and Geographic Proximity”, *Weltwirtschaftliches Archiv*, 132, pp. 97-115.

StataCorp (1999) *Stata Statistical Software: Release 6.0*, Stata Corporation, College Station.

Trajtenberg, M. (1999) “A Penny for Your Quotes: Patent Citations and the Value of Innovations”, *RAND Journal of Economics* 21 (1), pp. 172- 87.

Verspagen, B. (1997) “Measuring Intersectoral Technology Spillovers: Estimates from the European and US Patent Office Databases”, *Economic Systems Research* 9(1), pp. 47-65.

Verspagen, B., T. van Moergastel and M. Slabbers (1994) *MERIT concordance table: IPC - ISIC (rev.2)*, MERIT Research Memorandum, 94-004, Maastricht.

Appendix 1. Density and survival functions for the Pareto and the exponential-Weibull cases.

I derive the density and survival functions with the explanatory variables suppressed. If initial returns are of the Pareto type the density function is given by:

A1) $f(R_0) = aR^{a} R_0^{-(a+1)}$

In A1), the parameter α is given from the distribution function while the variable R is the lower constraint for initial returns (parameterised to give a survival rate equal to 1 at age 0). S(t) then becomes equal to the conditional probability that $R_0 \geq z$ given that $R_0 > C_0$. In this case the survival function, i.e. the probability of a patent surviving at least to age t, is given by:

A2) $S(t) = R^{a}\left(C_t e^{dt}\right)^{-a}$

Disregarding as in the text the term C'(t), this gives the density function of patent age and the associated hazard rate equal to

A3) $f(t) = daR^{a}\left(C_t e^{dt}\right)^{-a}$

A4) $h(t) = da$

In the Pareto case therefore, the hazard rate becomes constant and the survival function can be estimated as exponential.

If initial returns are exponential-Weibull distributed, the density function is given by:

A5) $f(R_0) = AaR^{-a} R_0^{a-1} e^{-r^{a} R^{-a}}$

In A5), the parameters α and R are the shape and scale parameters, respectively. A is a scaling parameter set equal to $\exp(C_0^{\alpha} R^{-\alpha})$. This is used to define S(t) so as it becomes equal to the conditional probability that a patent survives at least to t, given that $R_0 \geq C_0$. Now, define $C_t = KK_t$ where $K_0 = 1$. In this case the survival function is given by:

$$\text{A6)} \quad S(t) = e^{-K^{a} R^{-a} \left(K_t^{a} e^{adt} - 1\right)}$$

This is similar to a Gompertz function for survival time and, when the term C'(t) is ignored, the implied density and hazard rate functions become:

$$\text{A7)} \quad f(t) = K^{a} K_t^{a} R^{-a} ade^{\left(adt + -K^{a} R^{-a} \left(K_t^{a} e^{adt} - 1\right)\right)}$$

$$\text{A8)} \quad h(t) = K^{a} K_t^{a} R^{-a} ade^{adt}$$

Appendix 2. Weights used in non-linear least squares

The non-linear regression model presented in equation 11) has the composite error term u_t. This term consists of errors from the renewal rule, w, assumed to have zero mean and constant variance and the binomial sampling error v_t which is heteroscedastic. The true dependent variable is $y_t = \Phi^{-1}(1 - S_t)$ but I only observe the empirical Kaplan-Meier estimate $S^*_t = S_t + v_t$. Following Amemiya (1981), I substitute this into y_t and approximate around S^*_t to get:

$$\text{A9)} \quad \Phi^{-1}(1 - S^*_t) = H + \frac{\partial \Phi^{-1}}{\partial S_t}\Big|_{S_t} (S^*_t - S_t)$$

$$= H + \frac{1}{\left[j\left(\Phi^{-1}(1 - S_t)\right)\right]^2} (S^*_t - S_t)$$

$$= H + v_t$$

where $$H = \frac{-m + \ln C_t + dt}{s}$$

We have $Ev_t = 0$ and

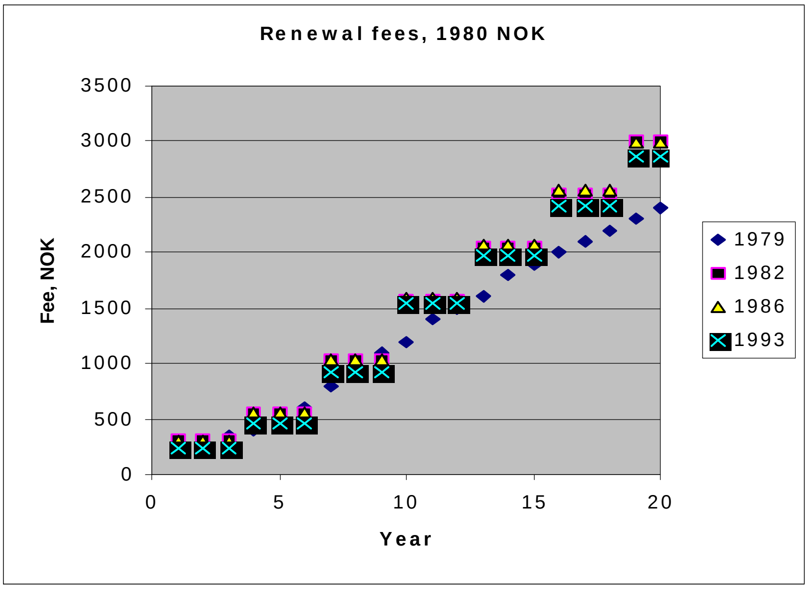
A10) $$Vv_t = \frac{S_t(1-S_t)}{n\left[j\left(\Phi^{-1}(1-S_t)\right)\right]^2}$$

Therefore, equation 11) is a non-linear regression model with a heteroscedastic error term. In order to construct the appropriate weights, I follow the procedure described in Schankerman (1998). I first estimate the model by unweighted least squares. This gives estimates of u_t. Together with v_t this allows me to construct w and w^2 which are used to construct σ^2_u serving as weights in the non-linear least squares regression.

Appendix 3. Data construction and renewal costs

The data on renewal fees are from Patentstyret (1979, 1982, 1986 and 1993). The Norwegian renewal costs change from time to time, but in such a way that real costs for patents of a certain age are almost constant. The fees are increasing stepwisely in age. After 1983, fees are constant for periods of three years. Figure A1 graphs the Norwegian renewal fees in real 1980 NOK.

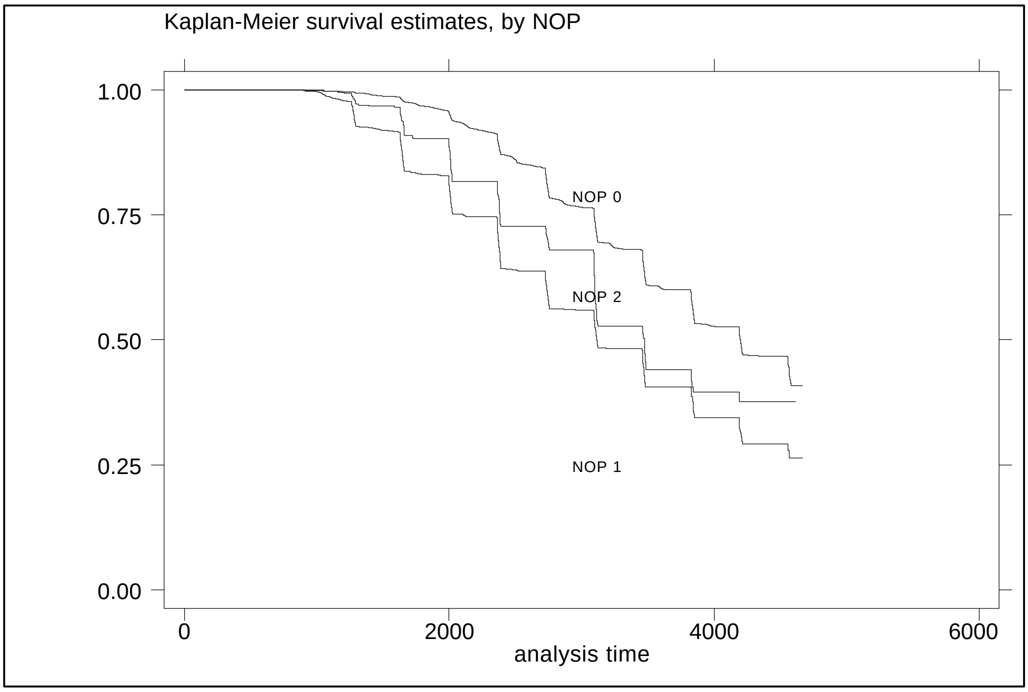
Figure A1



Source: Patentstyret (various years).

In the figure, I have included the 1993-renewal fee scheme in order to demonstrate that expectations of a constant real fee scheme are in some sense 'rational'. To simplify analysis, the average of the three relevant renewal fee schemes (1979, 1982 and 1986) was calculated, converted to daily rates and used in the analysis. A simple regression of the stepwise log of daily costs on age (in terms of days) gave a constant equal to –0.6 and the coefficient 0.00054 for days (R^2=0.85).

Figure 1.



Note: NOP=0 is for patents applied for in Norway by foreigners
NOP=1 is for patents applied for in Norway by Norwegians and not identifiable in the EPO database.
NOP=2 is for patents applied for in Norway and identifiable in the EPO database.

Figure 2. Histogram of patent citations

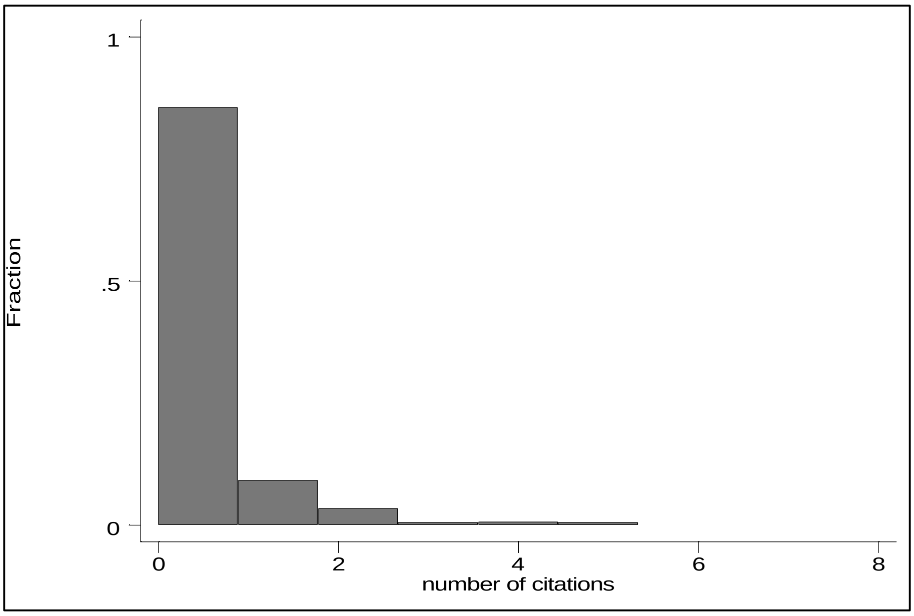
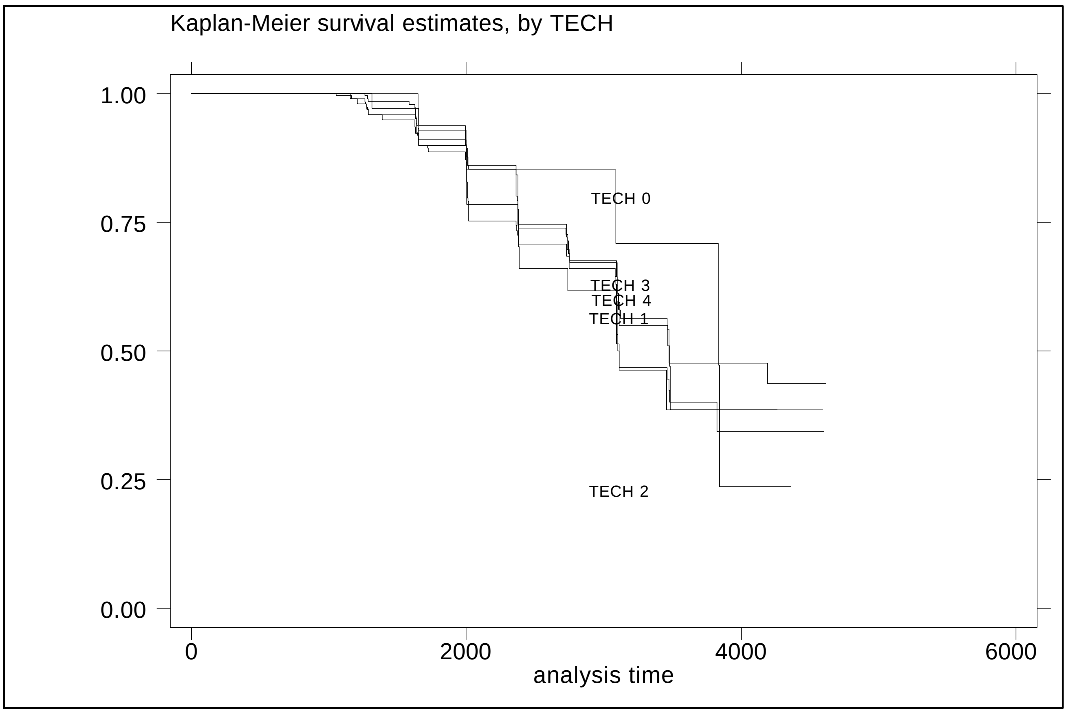


Figure 3.



Note: TECH=0 is for a composite group that could not be associated with R&D intensity. TECH=1,2,3 and 4 are for groups with increasing R&D intensities.

Figure 4.

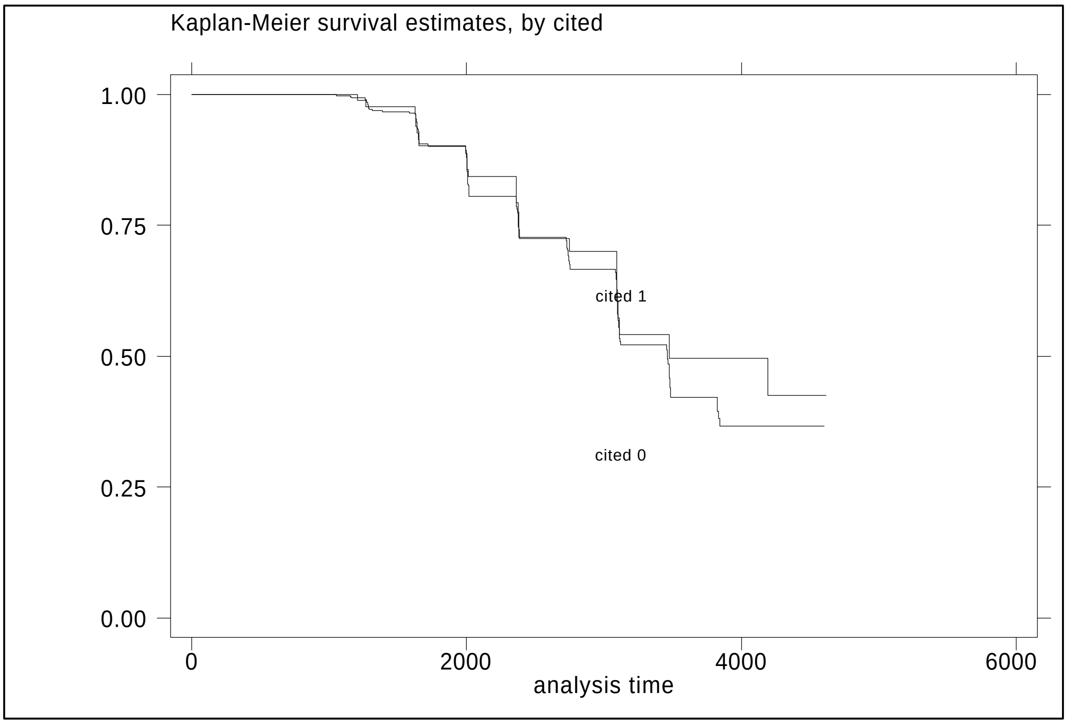


Figure 5

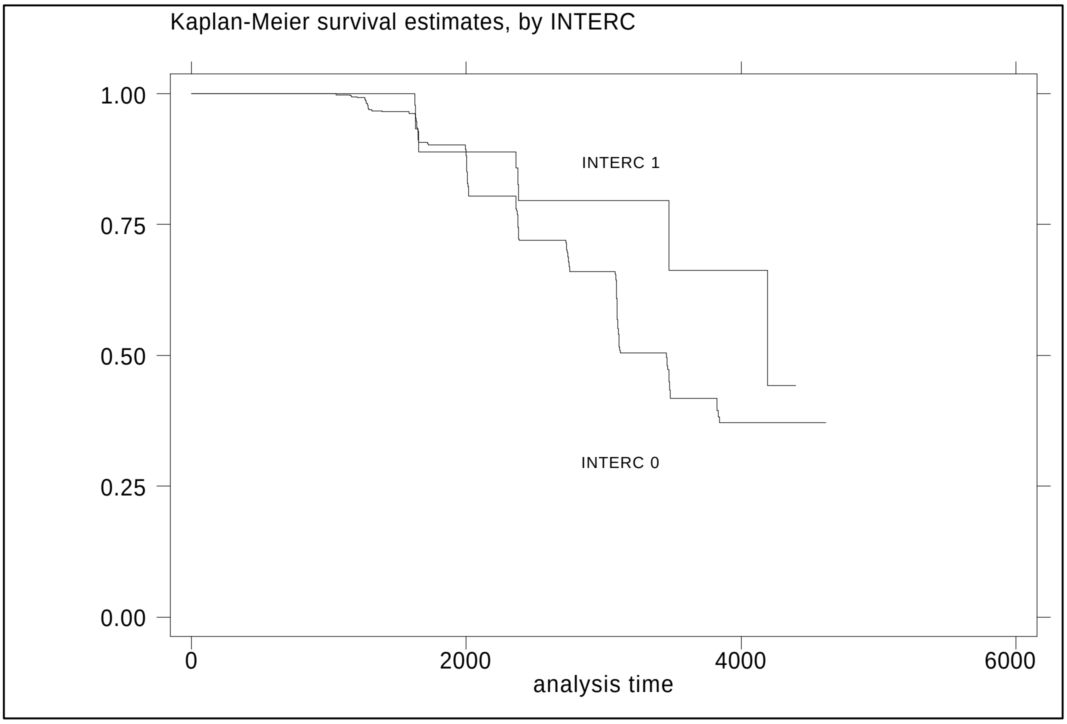


Figure 6

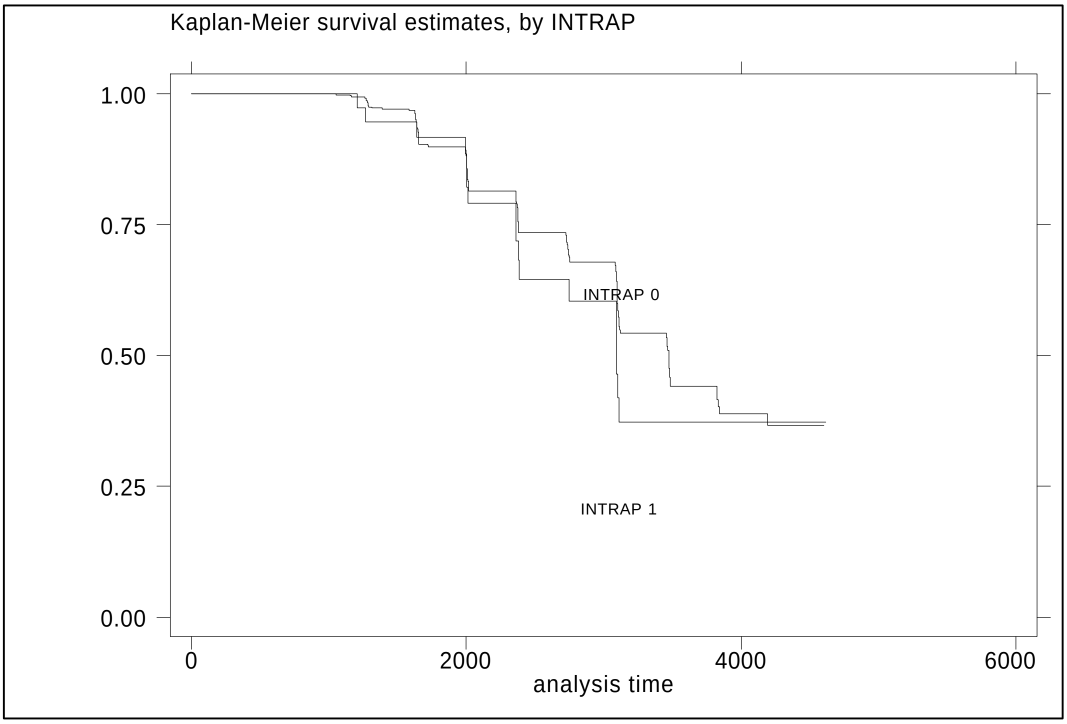


Table 1. Tests for equality of survival functions (significance of chi2 test)

	Log rank	Wilcoxon
Cited versus not cited	0.57	0.71
Intra-IPC cites versus no cites	0.41	0.38
Inter-IPC cites versus other patents	0.13	0.21

Note: The difference between the log rank test and the Wilcoxon test is that the Wilcoxon test statistic weights the difference between observed and expected number of failures (to pay renewal fee) by the total number of observations at risk. This is appropriate when the hazard functions vary non-proportionally.

Table 2. Cox model for survival of patents, estimated proportional hazard rates (p-values in parentheses)

N=598 **Failures: 163**	**Common model.**	**Accounting for self-citations.**
Log likelihood	-862.47	-861.05
CITES	0.17 (0.000)	0.12 (0.000)
INTRA	6.59 (0.000)	8.06 (0.000)
Self-citations	-	1.23 (0.820)
INTRA self citations	-	1.15 (0.906)
High-tech	1.41 (0.493)	1.42 (0.473)
High-medium tech	1.50 (0.392)	1.50 (0.385)
Low-medium tech	1.75 (0.235)	1.76 (0.231)
Low	1.73 (0.300)	1.72 (0.309)
1981	1.07 (0.821)	1.06 (0.838)
1982	1.30 (0.404)	1.30 (0.405)
1983	0.82 (0.557)	0.83 (0.566)
1984	1.95 (0.026)	1.97 (0.024)
1985	1.09 (0.817)	1.05 (0.885)
1986	1.89 (0.062)	1.91 (0.057)
1987	1.02 (0.961)	0.97 (0.947)
1988	0.65 (0.581)	0.66 (0.584)

Table 3. Regression results – value distribution of patents (p-values in paranthesis)

Lognormal distribution	N=563 adj. R^2=0.90
ì	5.45 (0.000)
ä	0.0012 (0.000)
ó	1.79 (0.000)

Note: Log-normal distribution estimated by weighted non-linear squares. The number of observations for the log-normal distribution is less than the number of observations because the survival function was equal to one for 38 observations.

Table 4. Implied quantiles for log-normal value distribution, 1980 NOK, (yearly return at age 0)

Quantile	**Return**
0.25	25 985
0.50	86 496
0.75	290 407
0.95	1 601 427
Mean	438 374

Note: The table is constructed by making 50 000 pseudo-random draws from a normal distribution with parameters taken from table 3, exponentiating the resulting numbers (in order to obtain a log-normal distribution) and multiplying them with 365 (in order to obtain yearly rates of return). See footnote 21. The resulting deviations from the exact distribution are the result of pseudo-random draws.

Table 5. Estimation of survival of patents, log-normal distribution of value (p-values in parentheses)

N=598 Failures: 163	Common model	Accounting for self-citations
Log likelihood	-2665.0	-2664.0
CITES	1154 (0.002)	1336 (0.001)
INTRA	-1251 (0.001)	-1360 (0.000)
Self-citations	-	-438 (0.670)
INTRA self citations	-	169 (0.895)
High-tech	-406 (0.473)	-394 (0.485)
High-medium tech	-443 (0.411)	-445 (0.409)
Low-medium tech	-623 (0.246)	-623 (0.244)
Low	-581 (0.344)	-565 (0.362)
1981	-384 (0.299)	-365 (0.322)
1982	-778 (0.044)	-786 (0.041)
1983	-248 (0.516)	-252 (0.507)
1984	-1150 (0.001)	-1165 (0.001)
1985	-557 (0.136)	-510 (0.173)
1986	-1190 (0.001)	-1206 (0.001)
1987	-673 (0.131)	-642 (0.152)
1988	-58 (0.918)	-63 (0.911)
Constant	5515 (0.000)	5504 (0.000)
Sigma	1575 (0.000)	1570 (0.000)

Figure 7.