Under Threat: Potential Competition, Litigation and the Private Value of Patent Protection

Jean Olson Lanjouw
Yale University and
London School of Economics and Political Science

The Toyota Centre
Suntory and Toyota International Centres for
Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel.: (020) 7955 6674

Discussion Paper
No. EI/6
July 1992

This paper has benefited from useful discussions with Stephen Howes, Peter Lanjouw, Ariel Pakes, Mark Schankerman and John Sutton as well as comments from participants in the LSE graduate seminar. All remaining errors are my own.

Abstract

This paper investigates the effect of the threat and occurrence of patent litigation on the private value of patent protection. Potential challenges are introduced into a renewal model as a factor in patentee decisions as to whether a patent is worth maintaining. The model yields testable predictions about renewal probabilities. Data for post WWII German patents support the hypothesis that the need to defend patent rights influences patentee behaviour. The paper concludes with a discussion of how this factor may be incorporated in patent value estimations and what the results imply for the interpretation of patent data.

Keywords: Patents; patent litigation; patent value estimations; litigation; patentee behaviour.

1. Introduction

Patent renewal data has been used to estimate distributions of the private value of patent protection (Schankerman and Pakes, 1986; Pakes, 1986; Schankerman, 1991). In these models, a patent renewal rule is derived based on annual renewal fees, the implied hazard probabilities are calculated and then fit to observed hazard proportions. Although allowance is made for depreciation in the returns to patent protection due to subsequent innovation and imitation, events which lead to the attenuation of patent rights are not modelled explicitly - challenged and unchallenged patents are described by the same hazard probabilities. However, inspection of German data (described below) reveals that patents of the same age have strikingly different hazard proportions depending on whether or not they have been recently granted. The pattern of hazard rate differences can be explained in a straightforward way as the outcome of patent litigation.

The threat of potential patent challenges has be analyzed using bargaining models which seek to explain the likelihood and level of out-of-court settlements between parties in conflict who wish to avoid litigation (ie, Meurer, 1989). This paper investigates the effect of the threat and occurrence of patent litigation on renewal behavior and the value of patent protection. In the most obvious sense potential competition for the use of an innovation is crucial to patent value - if no one would use the innovation in the absence of patent protection that protection is of little value. However, potential competitors also bring into question the validity and enforceability of patent rights. They pose the threat of infringements which may be costly to prosecute. While the first consideration is important because it affects the interpretation of estimated patent values, the second alters the form of the model which is appropriate to use in obtaining such estimates.

The empirical focus is on differences in renewal behavior between challenged and unchallenged patents. Section two presents an outline of the model. In section three, it is asked whether the two groups differ in any way relevant to renewal behavior *aside* from the fact that patents in the one group were the subject of a challenge. One practical problem, discussed in section three, is that challenges are not directly observed and must be inferred from the date of granting. In sections four and five a set of predictions about renewal behavior are derived from a model which includes challenge considerations. Discrimination between the models consists of showing that the occurrence of challenges does lead to significant differences in renewal behavior and that the observed differences are consistent with those predicted when challenge considerations are incorporated into the renewal model.

The data contains grant and renewal information for a sample of over 20,000 German patents, 1953-1988. Over the period the German Patent Office used two granting procedures. Before October, 1968, examination was automatically initiated after application. Annual renewal fees were due along with a publication fee only after a successful examination. Currently, applicants have 7 years to request and pay for a full examination. Annual renewal fees are due regardless of whether an examination has yet taken place. If a patent is not renewed then it lapses permanently. Details of the data set and a summary of the two regimes, fees and legal protection are in Appendices I and II.

2. The Model

A patentee is assumed to be a profit maximizer who will renew his patent as long as his expected return to patent protection, given the information at the time of the decision, is at least as high as the renewal fee. The expected return consists of the returns to protection in the current year, r, which are known, plus the expected value of maintaining the option of continued protection up to the maximum term.

A potential user of an innovation has three strategies. He may decide not to challenge. He may challenge the validity of the patent application using the patent office opposition procedures - an opposition challenge[1]. Alternatively, a competitor may simply use the innovation - an infringement challenge - and seek to have the patent revoked if prosecuted. If prosecuted for an infringement when the competing product or process is similar but not exactly the same as the patented innovation, defense may also be based on the claim that the act is not actually an infringement. Then the challenge does not concern the validity of the patent but rather the scope of the protection offered.

In the case of an infringement, in addition to full compensation for lost profits, the losing party must pay court costs as well as patent attorney and patent agent fees for both sides. These costs are related to the size of damages by a statutory fee schedule. In the case of an opposition challenge, these fees are substantially lower and neither party receives compensation (Korner, 1984).

---

[1]In recent years the annual number of patents undergoing an opposition challenge has been approximately 10-15% of the number finally granted (German Patent Office, 1987).

The annual profits available to the patentee having a monopoly over the use of an innovation[2] are denoted by the vector over ages π^m. In the absence of patent protection, if j competitors produce using the innovation in addition to the patentee, the annual profits earned by each agent including the patentee are equal and are denoted by the vector $\pi^{(j+1)}$, where $\pi^{(j+1)} \leq \pi^{(j)}$ all j. Agents are assumed to have the same cost functions. The total number of competitors who would use the innovation in the absence of patent protection is N. Each agent's profits when all produce using the innovation are denoted π^3 $\{\pi\}$ is used to denote the full vector set of profit opportunities over age $\{\pi^m, \pi^{(2)},...,\pi\}$. In any age, the returns to protection, r, is the difference π^m - π. Because π falls in N, the returns to the monopoly rights created by the patent increase in the level of competition. An innovation may be useful in several markets, where markets are defined by different sets of competitors, in which case the total returns to protection is the sum of returns across markets.

A formalization of the model is in Appendix III.

3. Are Challenged Patents Different?

Before turning to the question of whether the *event* of being challenged influences

---

[2]Profits earned from a monopoly over an innovation are not, in general, equal to industry monopoly profits. In the first, the production costs associated with the next best production technology available to producers who do not have rights to the innovation put a ceiling on price. The two would, however, be equal in the case of a new product innovation or a process innovation which is 'drastic' as defined by Arrow (1962).

[3]Firms are treated here as having no alternate forms of dissuading entry. This allows a one-to-one correspondence between the number of users in the absence of patent protection, N, and the level of competition. In general, N is the result of optimizing decisions on the part of firms regarding deterrence expenditure and will depend upon their ability to maintain positive profit in the face of entry and the cost of preventing entry. In this case, π is net of per firm deterrence costs. An increase in the level of competition, implying an increase in the cost of dissuading entry, again increases the patent returns to protection although it might not lead to a change in N.

renewal behavior, this section considers whether the two groups of patents, challenged and unchallenged, differ in any other relevant respect. In particular, are challenged patents of higher or lower value or can they be treated as coming from the same underlying value distribution?

Consider a situation with a given number of potential users, N. An innovation which has a proportionately higher set of profit opportunities $\{\pi\}$ than another enjoys higher returns to patent protection. The benefit to a challenger who successfully prevents a patent application from proceeding to grant or who has a patent revoked is the improved profits from using the innovation, π, after the decision. In the case of an infringement, the challenger also obtains profits over the period[4] of litigation, $\pi^{(2)}$. Therefore, given N, the increase in profit opportunities has a direct positive effect on the likelihood of challenge (see Appendix III, equations A.3, A.6, A.9).

Proposition 1: *Conditional on the level of potential competition, a positive relationship exists between returns to patent protection and the probability of challenge.*

On the other hand,

Proposition 2: *Conditional on the set of profit opportunities $\{\pi\}$, there is a negative relationship between the returns to patent protection and the probability of challenge.*

---

[4]While a patentee has the right to request an injunction, they are rarely granted and often not sought as they put the patentee at risk of having to pay damages to his competitor if infringement is not upheld (see Sperber, 1980).

The intuition for proposition 2 is that, a larger number of potential users[5] in the market increases the returns to protection. At the same time, the presence of many other potential users lowers the payoff to a challenger for breaking a patent because the innovation becomes available to all agents. In addition, the challenger pays the legal fees and, in the case of infringement, bears all of the risk of an unsuccessful challenge (equations A.3, A.6, A.9).

Thus there is a theoretical basis for associating high returns to patent protection with either a high or low probability of challenge. One must look to empirical evidence as to whether challenged patents can be treated as having the same distributions of returns as unchallenged patents. Since the distribution of returns determines the probability of renewal in each age, this translates into a null hypothesis that the hazard probabilities for the two groups are the same in each age[6].

The data do not indicate which patents have been challenged. However, once a patent application has been successfully examined, the only reason for a delay in granting is that someone has initiated an opposition challenge. As a result, grant age may be used as an indicator of opposition challenges. Under the early regime, patent applicants had no decision to make regarding the speed of the granting procedure. It may therefore be assumed that

[5]The number of potential users of an innovation may not be independent of its profit opportunities since both might be expected to be positively related to market size.

[6]Since hazard proportions are only informative about discrete segments of a distribution, rejecting equality of hazards is equivalent to rejecting equality of the distributions, but not vice versa. That is, the same hazard proportions can be generated by different distributions of returns. This drawback is not crucial here, first because the length of time and cost variation in the data increase the ability to discriminate using hazard proportions. Second, since hazard proportions are the fundamental piece of data, it is similarities and differences at that level which are relevant.

patents granted longer after application are more likely to have been through a challenge than those granted soon after application. This follows from the fact that one of the targets of the patent office is to avoid having any applications suffer lengthy delays in the examination procedure.

Under the current regime, the same positive relationship between grant age and the probability of having been challenged should hold. With an examination request at the latest age of 7, plus 2-3 years to be examined, the latest that a patent under normal circumstances would be granted is around age 11. Later granting indicates that an opposition was launched. Any patent granted very quickly is likely *not* to have been challenged. The group of patents granted in middle ages includes those who requested the examination early but were challenged and those who requested the examination late but were not challenged[7]. Nevertheless, in considering the whole span of grant ages, one may reasonably expect the proportion having been challenged to increase in grant age[8].

---

[7]In the procedures since January 1, 1981, grant occurs immediately following a successful examination without an opposition period. Consequently challenge need not cause a delay in granting. For this reason, patents granted after 1980 are not included in the empirical work presented here.

[8]Before using grant age as a proxy in the following tests, it is important to consider whether grant age may be related to returns independent of the occurrence of challenge. Under the early regime the answer is clearly no as the examination period is not related to economic value and patentees had no choice as to timing. Under the current regime, applicants do have a choice and data on average grant ages shows that a substantial percentage of patent applicants request the full examination before it is required (Lanjouw, 1992, Table 1.5). There are a number of possible reasons. The applicant may wish to use the results in support of patent applications in other countries or want a firm legal basis for licensing, pointing toward higher valued patents making early requests. On the other hand, early payment of the high examination cost can be used as a commitment on the part of the patentee to continue to granting and hence may delay challenges. It is a credible commitment because once the examination cost is paid it is sunk and the benefits of continuing to granting are enhanced. For patent applicants expecting low returns, this difference may be crucial to their request of the examination at all.

If the group of challenged and unchallenged patents derive from the same value distribution, then they have the same hazard probabilities over age. Using grant age as a proxy for the occurrence of a challenge, this translates into a test of whether patents granted at different ages have the same hazard probabilities. Vector[9] chi-square tests of the equality of hazard proportions across patents granted at different ages could not reject the null for either regime (old regime: p-value = .06, df 76; current regime: p-value > .99, df 137). The evidence indicates that opposition challenges, which delay granting, are not being launched against patents which generate a differing level of returns. Therefore, the maintained assumption will be that the only systematic difference between the two groups is the actual occurrence of challenge.

4. Post-grant Behavior

Introducing challenge considerations into a renewal model yields a straightforward prediction about how the behavior of patents just after granting is altered by having been challenged. Two other factors which influence post-grant hazard rates are also considered in order to disentangle their effect from that of challenges. The data is examined for evidence of the three factors by calculating, for each age, the post-grant hazard proportion, where 'post-grant' refers to the hazard rate out of patents granted in the previous age, and the distant-grant hazard, where 'distant-grant' refers to the hazard rate out of patents granted two or more years earlier. The data is presented in the following section.

---

[9]The tests use 4-5 cohort, 3 grant age and 4-5 non-renewal age groups to avoid small sample sizes. The tests are also conditioned on country/technology categories. The impact of post-grant effects (see below) is removed by including only patents which survive at least two years past granting.

As evidenced by late grant ages, opposition challenges often extend for numbers of years. Some patents undergoing a challenge will begin to yield negative net annual returns at some point during the challenge period. If not in the midst of a challenge, the patentees would stop renewing. However, because obtaining the positive returns at the beginning of a challenge period is contingent on winning the challenge (i.e. through patent office compensation requirements or contingent royalty payments) the patentees continue to granting. At this point, they cease renewing. As a result, the post-grant hazard for challenged patents includes those patents which would have been dropped in that age regardless of having been challenged *plus* those which would have been dropped earlier if not for the need to continue through the challenge period in order to gain benefits contingent on winning.

Proposition 3: (Challenge Effect) *The post-grant hazard probability of a group of challenged patents will be greater or equal the distant-grant probability for the same age of a similar group of patents*[10].

In particular, this proposition holds when, as indicated by the chi-square tests, the return characteristics of the two groups are *identical*. Before turning to the data for evidence of challenge effects in the post-grant hazard rates, two other factors which may alter these hazard rates are presented, lumpsum effects and changes in scope.

---

[10]At a given age, the post-grant hazard of challenged patents will also be greater or equal the post-grant hazard of unchallenged patents. However, since challenge can only be inferred from grant age in this data, it is not possible to look for evidence on this point.

When litigation over infringements arises after granting, the challenge effect would increase the hazard probability for challenged patents in the age following the resolution of their suits.

Costs related to granting are quite high in Germany relative to renewal fees in the early years. For this reason they can exercise a selection effect on the hazard rates of recently granted patents. Consider a simple example: a group of patents generates initial returns r_1, distributed $f(r_1)$, which decay at a known constant annual rate, δ (so $r_a = r_1 e^{-\delta(a-1)}$). Note that because returns are declining in age, if current net returns are zero then they will never be positive in the future. This means that the minimum level of returns which will lead to renewal in age a is $r_a = c_a$, the renewal fee.

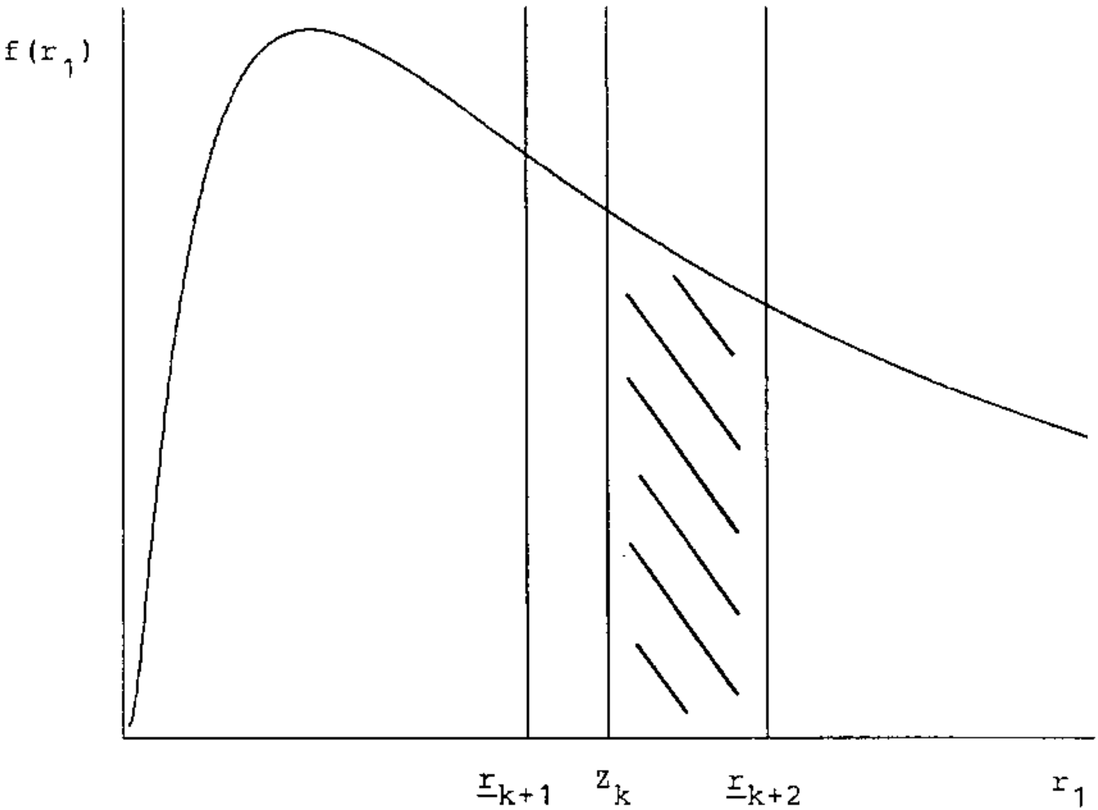
A patentee will request a full examination and pay the publication cost if and only if net returns over the examination period plus expected future returns contingent on being granted are greater than the examination costs. Suppose granting takes two years. Denote the minimum level of <u>initial</u> returns required for patents granted at age k to have requested an examination in k-2 as Z_k.

The implications for renewal behavior may be seen in Figure 1. The curve represents an initial distribution of returns. At any age a, the minimum level of <u>initial</u> returns which will lead to renewal is $\underline{r}_a = c_a/e^{-\delta(a-1)}$. If Z_k is at the level indicated in the figure, it is clear that there are no patents with $Z_k \leq r_1 < \underline{r}_{k+1}$ and consequently, *no* patents granted in age k will drop just after granting. Those in the shaded region between Z_k and $\underline{r}_{k+2}$ will drop in age k+2 and the observed hazard is that region divided by the area above Z_k. Proposition 4 follows.

Proposition 4: (Lumpsum Effect) *Ceteris paribus, as a result of high granting costs, the post-grant hazard probability for a given age is less than or equal the distant-grant hazard probability for the same age.*

Figure 1

Effect of Examination Costs



A final influence on the value of protection which may alter post-grant renewal behavior concerns the scope of that protection. Learning about the scope of protection during the granting procedure may occur due to the full examination or due to opposition proceedings. The coverage of the granted patent may be more circumscribed than that of the application. (Claims may not be added so the coverage cannot be broader.) Because of this, there may be some uncertainty on the part of a patent applicant about just what protection will be offered by the patent previous to granting.

As with uncertainty about granting and winning challenges, *ex-ante* uncertainty about the breadth of protection increases the minimum level of returns required to request an examination, where returns refers to those created by a patent granted the rights delineated in the original application. As a result, uncertainty about scope reinforces the lumpsum effect for all patents.

Some patentees learn that their claims will be given a narrow interpretation. Suppose a patentee believes himself to be the subject of an infringement where the challenger is using a product or process similar to the patented innovation. The challenger may defend his action by attempting to revoke the patent or he may defend his action by claiming that he is not actually infringing on the rights covered by the patent specifications. If the patent court upholds the validity of the patent but rules that the challenger did not infringe, the patent remains but with narrower bounds than the patentee had anticipated when suing for infringement. Similarly a patentee may learn that scope will be restricted during an opposition challenge. In both cases, some patentees may decide to let their patents lapse in light of the new information, thus reinforcing the challenge effect.

Proposition 5: (Scope) *Uncertainty about the breadth of protection increases the lumpsum effect while learning about the breadth of protection contributes to the challenge effect in magnifying post-grant hazards.*

5. Predicted Post-grant Effects and Empirical Evidence

This paper has put forward a number of factors which may cause hazard rates for newly granted patents to differ from those of patents granted earlier. The challenge effect (Proposition 3) is caused by a cumulation of patents delaying non-renewal until after granting because, due to having been challenged, granting is necessary to gain benefits in the form of compensation. This leads to a relatively large proportion of patents dropping just after grant. On the other hand, the lumpsum effect (Proposition 4) suggests that patentees willing to pay high examination costs should be willing to pay renewal fees for the early ages, lowering post-grant hazard rates. Considerations of scope (Proposition 5) would tend to magnify both of the above factors.

Predictions about the direction and strength of these three effects are summarized in Table 1. -/+/0 indicate that the post-grant hazard is predicted to be lower/higher/equal compared with the distant-grant hazard. Double symbols represent strength but only relative to the *same* factor in different ages. For example, -- for lumpsum in 'Very Early' ages does not mean that the effect of this factor is more important than the scope effect. It merely suggests that high granting costs are likely to be relatively more important at early ages than at later ages. 'Very Early' are those ages where there is little chance that a patent granted in (age-1) has been challenged. Hence, the *only* reason for dropping is obsolescence, a drastic fall in

Table 1

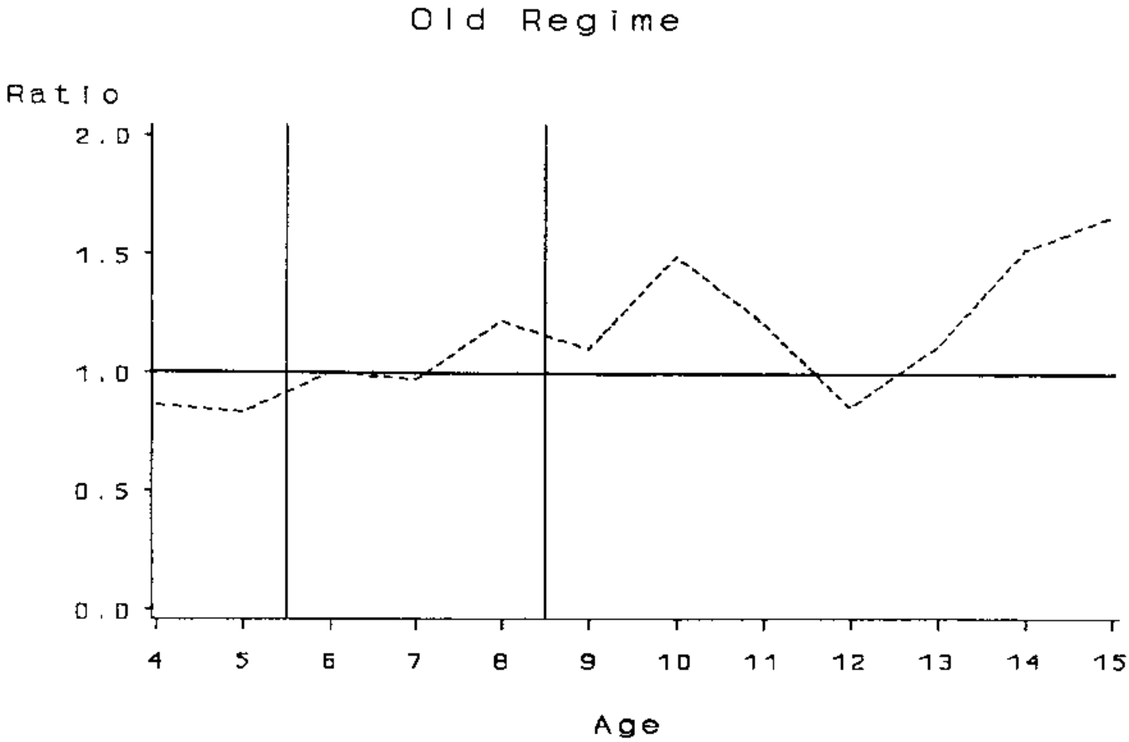
Predicted Effects on the Ratio of Post-grant to Distant-grant Hazards

	Ages		
	Very Early	Middle	Late
Regime			
Early	3-5	6-8	9+
Current	3-6	7-12	13+
Factor			
Challenge	0	+	++
Lumpsum	--	-	0
Scope	-	0	+

returns. 'Late' includes ages for which a grant almost surely has been challenged. 'Middle' ages are likely to include both challenged and unchallenged patents. Can the combination of these effects explain the differences between post-grant hazards and distant-grant hazards observed in the data? Taken together, the predictions suggested in Table 1 are that post-grant hazard proportions should be relatively low when granting occurs in the early ages, relatively high when granting occurs in late ages, with no clear prediction in the middle years.

The main empirical data regarding grant effects on renewal behavior are summarized in Figure 2, which displays the ratio of post-grant to distant-grant hazards. Details are in Table 2. The first column of the table contains post-grant hazards and the second column contains the average hazard for a given age taken over those patents granted two or more ages earlier

Figure 2
Ratio of Post-grant to Distant-grant Hazards



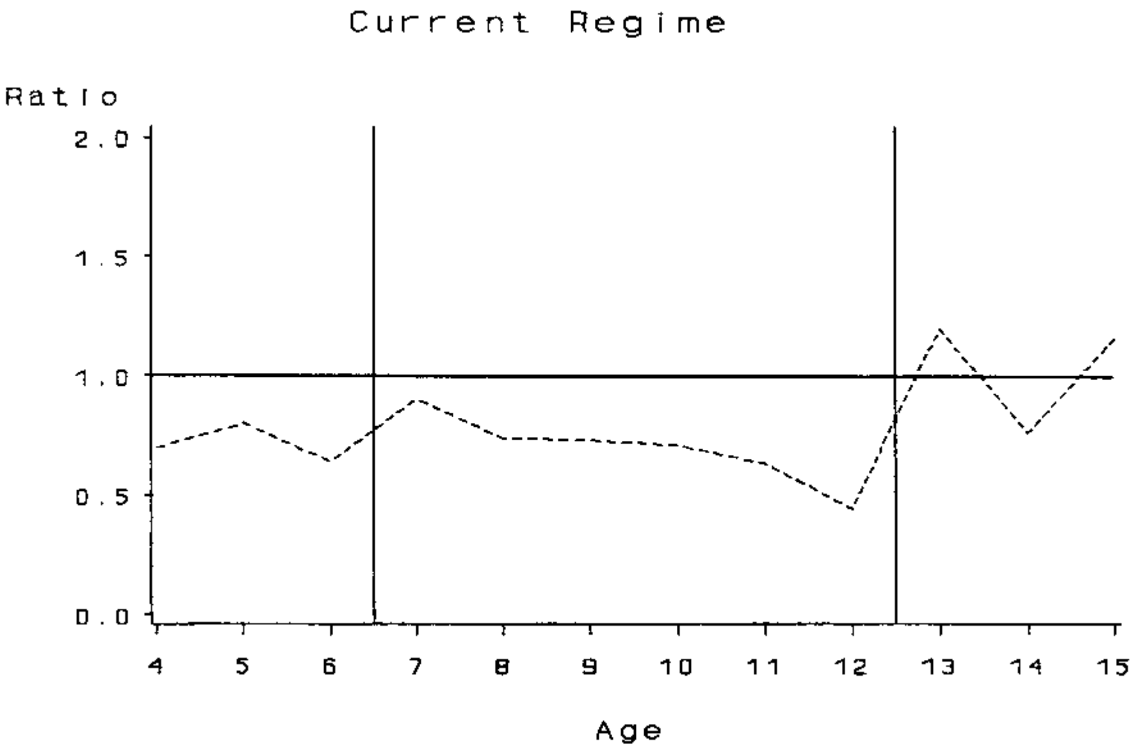


Table 2
Post-Grant vs Distant-Grant Hazards

Age	Post-grant Hazard[1] (1)	Distant-grant k=2 to (age-2)[2] (2)	Ratio (1/2) (3)	(Grants in k=age-1)/ Total in force (4)
Old Regime				
3	.045 (.006)			
4	.060 (.007)	.069 (.008)	.86 (.14)	.54
5	.076 (.008)	.091 (.007)	.83 (.10)	.32
6	.111 (.011)	.111 (.007)	1.00 (.12)	.20
7	.108 (.013)	.111 (.007)	.97 (.13)	.16
8	.147 (.017)	.120 (.007)	1.22 (.16)	.10
9	.157 (.020)	.142 (.009)	1.10 (.16)	.09
10	.201 (.025)	.135 (.008)	1.49 (.20)	.06
11	.160 (.023)	.132 (.010)	1.21 (.20)	.06
12	.153 (.029)	.179 (.013)	.85 (.17)	.04
13	.188 (.030)	.169 (.008)	1.11 (.18)	.02
14	.285 (.033)	.188 (.030)	1.52 (.19)	.02
15	.261 (.011)	.157 (.006)	1.66 (.10)	.02
Current Regime				
3	.055 (.012)			
4	.048 (.009)	.068 (.014)	.70 (.20)	.69
5	.060 (.008)	.075 (.008)	.80 (.14)	.45
6	.056 (.007)	.087 (.007)	.64 (.09)	.33
7	.083 (.010)	.092 (.006)	.90 (.12)	.22
8	.082 (.012)	.112 (.008)	.74 (.12)	.16
9	.099 (.015)	.136 (.011)	.73 (.13)	.16
10	.093 (.015)	.130 (.012)	.71 (.13)	.17
11	.105 (.020)	.167 (.012)	.63 (.14)	.11
12	.071 (.022)	.162 (.018)	.44 (.14)	.09
13	.254 (.037)	.214 (.016)	1.19 (.21)	.06
14	.125 (.034)	.164 (.022)	.76 (.22)	.05
15	.229 (.033)	.199 (.019)	1.15 (.20)	.03

Notes:

1) In columns 1 and 2 the hazard for a given age a is calculated in two stages. First an average over technology, country and cohort is obtained for each year, weighted by n_a. Second, the unweighted average over years is calculated.

2) Estimated standard errors are in parentheses. Because technology, country, cohort and grant age groups are independent, variances at each stage are calculated as sums of the variances of the constituent hazard estimates weighted by the relevant coefficients.

3) The variance of the ratio in column 3, (h_1/h_2), is calculated using a Taylor expansion:
$Var(h_1/h_2) \cong [1/(h_2)^2][Var(h_1)] + [h_1/(h_2)^2][Var(h_2)]$.

(distant-grant hazards).[11] For example, under the old regime, the proportion of patents granted in age 13 which do not renew in age 14 is .285 while the proportion of those patents granted in ages 2-12 which are renewed through age 13 and lapse in age 14 is .188. The third column, that shown in Figure 2, is the ratio of post-grant to distant-grant hazards. The fourth column gives the number of patents granted in age-1 as a ratio of the total number in force. Thus, a large deviation of the value in column three away from 1 combined with a high value in column four indicates that grant related factors have a large impact on the all inclusive aggregate hazard rates.

Recalling that the only reason for post-grant drops in the Very Early ages is obsolescence, one obtains some information about its magnitude from the table. The data reveal obsolescence rates of 4.5 to 7.6% in the Very Early ages. Apart from its independent interest, this should be borne in mind when interpreting the ratios in column three (and Figure 2). Age specific obsolescence contributes to post-grant and distant-grant hazards equally. In looking for evidence of the three effects summarized in Table 2, it is differences in the incremental hazards, over and above obsolescence, which is relevant. For example, suppose that there is an obsolescence rate of .05 in age 15. Subtracting .05 from the hazards in columns one and two, the ratio in column three would increase, under the old regime, from 1.66 to 1.97 and, under the current regime, from 1.15 to 1.21. In general, any positive obsolescence hides differences and brings the ratio towards 1.

---

[11]The hazards were calculated in two stages. First a weighted average over industry, country (and in column 2, cohort) groups was calculated for each year and then an unweighted average was taken over years. This ensures that years are given the same weight in both sets of hazards. The procedure was used to avoid biases due to the combination of a truncated sample and systematic year effects such as trend falls in the real renewal fee schedule.

Turning first to the old regime, the pattern of hazards is clearly in accordance with the predictions in Table 1. In the Very Early years the post-grant hazards are lower than the distant-grant hazards and the ratio is less than one. In the Middle years, where the predicted net effects are ambiguous the empirical net effects are also mainly zero. In the Late years post-grant hazards tend to be the same or higher indicating challenge effects. The results under the current regime are also supportive. They are somewhat less clear but this is not surprising in light of how long the Middle years stretch and the increased choice in the granting procedure. Here the lumpsum effect plays an important role, continuing to lower post-grant hazards throughout the Middle years. The fact that lumpsum costs increased substantially under the current regime from approximately 125 to over 400 1975 DM lends support to the identification of this effect with examination and publication costs[12]. In the Late ages, when predicted effects are all zero or positive, the post-grant hazards do increase dramatically relative to the distant-grant hazards with the ratio moving above one. Again, this fact suggests that challenges do influence post-grant renewal behavior.

6. Life with Litigation

From the preceding analysis one is able to draw a number of important conclusions bearing specifically on the appropriate formulation of patent renewal models. The conclusions also have broader implications, particularly regarding the interpretation of patent value and patent count data in relation to innovative output.

[12]The fact that the lumpsum effect bites (the ratio is far below 1) indicates that there are a substantial number of patentees making marginal decisions. This fact is crucially important to the ability to actually estimate renewal models successfully.

From the simple observation that the vectors of post-grant and distant-grant hazard proportions differ there is the presumption that grant proximity should be accounted for in a model of renewal decisions. This is particularly true in light of the fact that newly granted patents comprise a fairly large percentage of the total patents in force for ages up to 11 or 12 (see Table 2, column 4). Their inclusion in an undifferentiated aggregate hazard rate distorts the recorded pattern over ages in hazard proportions and dilutes the potential information content of both series.

One straightforward approach to this concern is to not include patents in the calculation of hazard proportions until they have renewed a number of years past granting. (Lanjouw, 1992. If disaggregated patent data is being used, the reduction in sample sizes may be a drawback.) At this point, their renewal behavior is no longer influenced by the factors discussed in the previous section. The one exception would be the case of an infringement challenge which occurs after granting and which could result in challenge or scope effects. Absent information on which patents have undergone an infringement challenge, there is not much one can do about this possibility aside from noting that such challenges are generally resolved informally[13] with only a small number of total patents ever the subject of litigation proceedings.

While subsetting the data to avoid post-grant effects goes some way to dealing with the issues raised in the preceding sections, the evidence indicating that litigation influences renewal decisions has more fundamental implications for the modelling of patent renewal.

---

[13]A study of patentees in Germany by the US Department of Energy (1982) found that large corporations rarely litigate and overall more than a third of suits filed are settled out of court.

Patent rights are granted but not actively enforced by the state. They have meaning only insofar as the patentee is willing to defend them - not a costless exercise. A patentee's decision to renew, in a competitive environment where potential infringers are poised to challenge his rights, is based not only the value of protection but on the costs of maintaining it. These costs depend on such factors as the level of legal fees, the time necessary to prosecute a suit and the probability of winning[14]. The need to be willing to defend against infringements enters the renewal decisions of all patentees, whether or not they are ever challenged. As a result, the minimum level of returns required to induce a patentee to renew at each age increases relative to that required under the assumption that patent rights are perfectly and costlessly enforced. This change in the renewal decision rule alters model hazard rates, which form the basis of patent value estimations.

Recognizing the threat of litigation also has implications for making the link between the value of patent protection and that of the underlying innovation. Not all innovations are patented and the threat of challenges may alter the value characteristics of those which are. For example, if the probability of a patentee winning a challenge, w, varies and the covariance between w and the value of innovation (π) is either zero or positive, one would see a skew to the right in the value distribution of that innovation which is patented relative to total innovation. Take the case of zero covariance. Independence implies that any two subsets of the total population of innovations, one with a high probability of winning challenges and the other with a low probability, will have the same expected distribution of

---

[14]With common knowledge across patentees and potential users, a patentee who would be unwilling to defend will be challenged with probability one. If potential users are uncertain as to whether a patentee would defend, the patentee's costs are lowered to the extent that this uncertainty dissuades potential challengers from acting.

$\{\pi\}$. However, the minimum level of returns which would entice a potential patentee to apply is lower for the first than for the second subset (equation A.16). Similarly for all subsets with successively higher probabilities of winning. The distribution of $\{\pi\}$ for those who patent among all subsets taken together is an aggregation of truncated distributions of $\{\pi\}$. Because the truncation point varies, the aggregate distribution of those innovations in a population which are patented is *not* a truncated version of the distribution of $\{\pi\}$ in the total population but is more skewed to the right.

These factors may also cause differences in the propensity to patent across populations and should be considered when drawing inferences from <u>numbers</u> of patent applications or grants. For a given w, the minimum level of returns necessary to induce an inventor to apply for a patent (or similarly to continue to granting) is a function of the number of potential users N. This establishes a cutoff point - any part of the value distribution of innovations below the cutoff point will not be in a count of applications (grants) (equation A.16). The observation that the number of applications (grants) has fallen over time may be explained in many ways as combinations of three forces at work: Changes in the value of innovation, changes in the number of innovations, and/or movement in the cutoff point due to changes in the level of competition. Changes in potential competition may be due to alterations in anti-trust policies, openness and ease of trade, advertising regulations and cost, etc. Changes in intellectual property policy which bear on the success of patentees in defending their rights would have similar influence on the cutoff point.

With the ever increasing computerization of patent databases it should become feasible to obtain renewal data for patents differentiated by whether challenged, how, how often, whether they were successfully defended and how long it took. Such information is available for Germany beginning in 1981. It would be interesting to see whether the results of the preceding section would be strengthened using cleanly identified groups. One would expect to see more pronounced challenge effects for patents who have undergone a particularly long period of litigation. Information on challenges would allow the formulation of fuller renewal models. In particular, one could incorporate examination costs explicitly for those patents not challenged, thus accommodating lumpsum effects, while subsetting the data only for those challenged. Knowing which patents were challenged and how long they were the subject of litigation also opens up the possibility of modelling challenge effects directly. At the very least, information on the probability of winning (for which one would expect the observed proportion to serve as a reasonable approximation[15]) could be included in the model, thus reducing the burden of estimation.

The value of patent protection is inherently related to the presence of potential users of the underlying innovations. The analysis in this paper suggests that the threat and occurrence of challenges to patent rights bears on patentee decisions as to whether patent protection is worth maintaining. Further analysis using more finely differentiated patent renewal data would be useful in understanding how important these considerations are in lowering the benefits of the patent system. Both the modelling of patentee decisions and the inferences drawn from patent renewal data should be made in light of the strategic environment that makes patent protection necessary.

---

[15]The true average probability of winning is not exactly the observed proportion which is conditional on challenge occurring. However, there is no clear direction of bias. With common knowledge, patents with a high probability of being found valid will not be challenged. Those with a low probability will not be defended.

Appendix I

Characteristics of the Data

Total Sample Size	20,235 patents
Percentage of Total Patents Granted (1953-1980)	3.5 %
Country Groups - by nationality of the owner	Western Europe United States Japan
Technoloyy Groups	Computers Textiles Engines Pharmaceuticals
Range of Years	
Western Europe and United States	1953-1988
Japan	1963-1988
Range of Cohorts	
Western Europe and United States	1953-1980
Japan	1963-1988
First Annual Fee due for age	3
Maximum Age	
1953-1976	18
1977-	20

Mean Sample Size[1]

	Western Europe	United States	Japan
Computers	102	86	67
Textiles	102	38*	34*
Engines	102	48*	40*
Pharmaceuticals	80	37*	40*

Note:

1) An asterisk indicates that the technology/country/cohort cell samples are equivalent to the entire population of granted patents for every cohort covered.

Appendix II

Patenting Procedure, Fees and Legal Rights

Stage	Fees[16]	Legal Rights
Old Regime - until October, 1968		
Application examination follows without request	Application fee 100DM	No rights to an injunction and compensation at the discretion of the patent office
Decision to Grant	Publication fee 100DM plus cumulated annual fees	Rights to claim an injunction and to full compensation.
3 month opposition period for opposition challenges		
Granting	Annual renewal fees 100DM age 3 to 2900DM age 20	
Current Regime		
Application Preliminary examination follows without request	Application fee 75DM	
First publication	Annual renewal fees (continuing) 75DM age 3 to 2200DM age 20	No rights to an injunction and compensation at the discretion of the patent office
Request for full examination (by age 7)	Examination fee 325DM	
Decision to grant	Publication fee 100DM	Rights to an injunction and to full compensation
3 month opposition period for opposition challenges		
Granting	Annual renewal fees	

[16]Figures are approximate levels of the indicated fee in 1975 Deutchmarks (=.95 1988 US dollars).

Appendix III

The Model

The model does not explicitly incorporate a time discount factor nor the effect of depreciation on expected future returns. It should be understood that all values are in present value terms and that stochastic variables are in expectation given the information available at age 1. Profits and costs may vary over age but subscripts will be suppressed to ease notation.

I. Payoffs to Challenge

This section characterizes payoffs to challenge for group of N potential users of an innovation if the patent application is expected to be defended. If defense is not expected, all potential users infringe at zero cost.

1. **Single Opposition Challenge**

Expected payoff if the challenger wins:

$$\Sigma^{T}_{p+1}\pi - p(lf) \qquad (A.1)$$

where p is expected period of litigation, Σ is a summation over ages through the statutory age limit, T[17], and lf are annual legal fees. This event occurs with probability (1-w).

Expected payoff if the challenger loses:

$$-p(lf) \qquad (A.2)$$

This event occurs with probability w.

Expected payoff from opposition challenge:

$$(1-w)\Sigma^{T}_{p+1}\pi - p(lf) \qquad (A.3)$$

2. **Single Infringement Challenge**

Expected payoff if the challenger wins:

$$\Sigma^{p}_{1}\pi^{(2)} + \Sigma^{T}_{p+1}\pi \qquad (A.4)$$

This event occurs with probability (1-w).

---

[17]The summation of benefits to the challenger due to challenge should, to be more precise, run from 1 to the age at which the patentee would have let the patent lapse in the absence of challenge, denoted s. However, by definition of age s, after s the net value of the rights to exclusive use of the innovation is negligible (does not justify defense). π, the value of non-exclusive use of the innovation, is even less and the difference in summand has no impact on the equation.

Expected payoff if the challenger loses:

$$\Sigma^{p}_{1}\,\pi^{(2)} - \Sigma^{p}_{1}(\pi^{m} - \pi^{(2)}) - 2p(LF) \qquad (A.5)$$

where the sum of lost profits plus legal expenses $\{\Sigma(\pi^{m}-\pi^{(2)}) + p(LF)\}$ are due to the patentee as compensation. LF represent the higher cost of litigating an infringement challenge. This event occurs with probability w.

Expected payoff to single infringement:

$$(1-w)\{\Sigma^{T}_{p+1}\pi + \Sigma^{p}_{1}\pi^{(2)}\} - w\,\Sigma^{p}_{1}(\pi^{m} - 2\pi^{(2)}) - 2w\,p(LF) \qquad (A.6)$$

If a single challenge is optimal, it will be via an infringement iff the difference between the expected payoffs from first infringement and an opposition challenge alone (equation A.6 - equation A.3):

$$(1-w)\Sigma^{p}_{1}\pi^{(2)} - w\,\Sigma^{p}_{1}\,(\pi^{m} - 2\pi^{(2)}) - p(2wLF - lf) \qquad (A.7)$$

is positive[18].

3. **Additional Challenges**

The fact that a challenge has been made ensures that an additional user need not incur the costs of a challenge to gain the possible benefits in p+1. Therefore, if more than one challenge occurs, it cannot be optimal for a potential user to challenge *without* infringement.

Expected payoff to an infringement if first challenge was an opposition challenge:

$$(1-w)\Sigma^{p}_{1}\pi^{(2)} - w\Sigma^{p}_{1}(\pi^{m} - 2\pi^{(2)}) - 2w\,p(LF). \qquad (A.8)$$

If an opposition challenge has been made, then (A.7) < 0. It is clear from a comparison of (A.7) with (A.8) that if an opposition challenge is made then it cannot be optimal for any challenges to occur via infringement. Thus, if one opposition challenge is made it is neither optimal for other agents to challenge via another opposition challenge or via infringement so it is an equilibrium position.

Expected payoff to the jth infringement if previous challenges via infringement:

$$(1-w)\{\Sigma^{p}_{1}\pi^{(j+1)}\} - w\Sigma^{p}_{1}[(\pi^{m}-\pi^{(j+1)})/j - \pi^{(j+1)}] - 2w\,p(LF) \qquad (A.9)$$

where responsibility for compensation to the patentee is divided equally between the

---

[18]Large amounts of fixed capital investment in the use of the innovation or start-up sales/marketing costs would mitigate in favor of opposition challenges by raising the cost of infringements. If the challenger loses, the investments, net of resale, are a direct cost. If he wins, then there may be an interest cost if an injunction is served. On the other hand, having made a commitment to the innovation early may give the challenger positive benefits in terms of competitive position later.

challengers. The probability that the patentee wins, w, is independent of the number of challenges but the patentee must pay constant legal fees and court costs to prosecute each additional infringement. The first term in equation (A.9), the benefit of winning, falls in j, while the direction of change in the second, the cost of compensation, is ambiguous in general. The rate at which π falls in j will determine the equilibrium number(s) of users willing to infringe, $j^* \leq N$, such that either equation (A.9) < 0 for j^*+1 or $j^*=N$. The total change in equation A.9 need not be monotonic in j so j^* need not be unique.

II. Payoffs to Defense

1. **Opposition Defense**

Expected payoff from successful defense against an opposition challenge:

$$\Sigma^p{}_1[\ \Sigma_m(\pi^m-\pi) - c] + \Sigma_m V_{p+1} - p(lf) \qquad (A.10)$$

where Σ_p and Σ_m represent summation over ages and markets and V_a is the expected value of the patent from a under the assumption that future renewal decisions are made optimally ($V_a \geq 0$). A successful defense occurs with probability w.

Expected payoff from an unsuccessful defense against an opposition challenge:

$$\Sigma^p{}_1[\Sigma_m(\pi^m-\pi)- c]- p(lf) \qquad (A.11)$$

This event occurs with probability (1-w).

Expected payoff to defense against an opposition challenge:

$$\Sigma^p{}_1[\Sigma_m(\pi^m-\pi) - c] + w\Sigma_m V_{p+1} - p(lf) \qquad (A.12)$$

2. **Infringement Defense**

Expected payoff from a successful defense against infringement:

$$\Sigma^p{}_1[\Sigma_m(\pi^m - \pi) - c] + \Sigma_m V_{p+1} \qquad (A.13)$$

This event occurs with probability w.

Expected payoff from an unsuccessful defence against j infringements in each market:

$$\Sigma^p{}_1[\Sigma_m(\pi^{j+1}-\pi) - c]\ - \Sigma_m j\ 2p(LF) \qquad (A.14)$$

where j may differ in each market. Each infringement suit is assumed to require additional

legal fees.[19] The choice not to defend against one of several infringements implies the loss of all returns. An unsuccessful defense occurs with probability (1-w).
Expected Payoff to defense against j infringements in each market:

$$\Sigma^{p}_{1}[\Sigma_{m}((1\text{-}w)\pi^{(j+1)}+w\pi^{m}-\pi) - c] + w\Sigma_{m}V_{p+1} - (1\text{-}w)\Sigma_{m}j2p(LF). \qquad (A.15)$$

The expected payoff to defense against an infringement falls in the number of infringements[20].

III. Equilibria

There are a number of challenge and defense equilibria. If the patentee will not defend against a single infringement challenge (which implies that he will also not prosecute $j>1$ infringements) then the only equilibrium is infringement by all potential users. If a single infringement would be prosecuted but no defense would be made to a court challenge then a single court challenge in conjunction with N infringements is a unique equilibrium. Thus if either equation A.12 or A.15 is negative for $j=1$ then the unique equilibrium includes N infringements and no defense. There are a number of possible equilibria when both equations A.12 and A.15 are positive for $j=1$. Suppose that equation A.12 is positive and A.15 becomes negative for j^d infringements where $1 < j^d \leq N$. The following are possible equilibria:

i. no challenge
ii. court challenge with defense
iii. j^* infringements with defense if $j^d > j^*$
iv. N infringements without defense if $j^d \leq N$.

Therefore, both decision criteria being positive for $j=1$ is a necessary requirement for a patentee to be willing to defend his application through to granting. If $1< j^d \leq i$ then there may be multiple equilibria, one of which entails no defense (iv), so both decision criteria being postive for $j=1$ is not sufficient to ensure defense. $j^d > i$ is the minimum sufficient criterion.

Incorporating the need to defend against challenges, the decision rule based on the minimium necessary to ensure defense is to renew iff:

$$\Sigma^{p}_{1}[\Sigma_{m}\, w(\pi^{m}-\pi) - c] + w\Sigma_{m}V_{p+1} - (1\text{-}w)\Sigma_{m}\, N2p(LF) \geq 0 \qquad (A.16)$$

---

[19]If one of the j infringers is significantly larger than the others, an alternate strategy would be to prosecute that infringement alone, in the hope that a successful outcome will cause the smaller infringers to desist.

[20]Dynamic consistency issues arise with respect to defense. For an analysis see Lanjouw, 1992, Chapter 4.

References

Arrow, Kenneth (1962) "Economic Welfare and the Allocation of Resources for Invention" In R. R. Nelson (ed.) The Rate and Direction of Inventive Activity: Economic and Social Factors. NBER Conference Series no.13. Princeton University Press: Princeton.

German Patent Office (1987) Annual Report 1987. Press and Public Relations Section. Munchen.

Korner, E. (1984) "German, Federal Republic of," in M. Mellor (ed.) International Patent Litigation. Washington, D.C.: The Bureau of National Affairs, Inc.

Lanjouw, Jean O. (1992) The Private Value of Patent Rights: A Dynamic Programming and Game Theoretic Analysis of West German Patent Renewal Data, 1953-1988. Ph.D thesis. London School of Economics.

Meurer, Michael (1989) "The Settlement of Patent Litigation", Rand Journal of Economics. Vol 20, no 1. pp 77-91.

Pakes, Ariel (1986) "Patents as options: Some estimates of the value of holding European patent stocks". Econometrica, 54, pp. 755-784.

Schankerman, Mark and A. Pakes (1986) "Estimates of the value of patent rights in European countries during the post-1950 period." The Economic Journal, 96, pp. 1052-1076.

Schankerman, M. (1991) "The Private Value of Patent Rights in France, 1969-87: An Empirical Study of Patent Renewal Data", NBER Working Paper, July.

Sperber, P. (1980) Intellectual Property Management. Law - Business - Strategy, vol. 3, Intellectual Property Series. 2nd edition. Clark Boardman Co., Ltd: New York.

US Department of Energy (1982) The Patent Systems of the Federal Republic of Germany and Japan: Personal Interviews with Patent Practitioners. Office of Technology Assessment. Washington, D.C.