NBER WORKING PAPER SERIES

PATENT PROTECTION: OF WHAT VALUE AND FOR HOW LONG?

Jean Olson Lanjouw

Working Paper No. 4475

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 1993

I would like to thank Mark Schankerman for many useful discussions throughout the course of this work. It has also benefitted from comments from Ricardo Curcio, James Davidson, Stephen Howes, Peter Lanjouw, Jan Magnus, Ashoka Mody, Ariel Pakes and Hugh Wills. Remaining errors are, of course, my own. This paper is part of NBER's research program in Productivity. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

NBER Working Paper #4475
September 1993

PATENT PROTECTION: OF WHAT VALUE AND FOR HOW LONG?

ABSTRACT

Empirical estimates of the private value of patent protection are found for four technology area - computers, textiles, combustion engines, and pharmaceuticals - using new patent data for West Germany, 1953-1988. Patentees must pay to keep their patents in force. A dynamic stochastic discrete choice model of optimal renewal decisions is developed incorporating both learning about an innovation and the market as well as the possibility of infringements. The evolution of the distribution of returns over the life of a group of patents is calculated for each technology using a minimum distance simulation estimator. Results indicate that learning is completed within 6 years, that obsolescence is rapid, and that the distributions of patent value are very skewed. Research and development (R&D) expenditures are calculated and patent protection as an implicit subsidy to investment in R&D discussed.

Jean Olson Lanjouw
Economics Department
Yale University
37 Hillhouse Avenue
New Haven, CT 06520
and NBER

1. Introduction

This paper provides estimates of the level and distribution of private value of patent rights to inventors in selected technology areas using West German[1] patent data. The results may be useful in addressing two different sets of questions. First, in many areas of economic research one would like a measure of innovation. Studies of the relative efficiency of government, private or academic R&D in the production of economically useful ideas; studies of spillovers or agglomeration effects on R&D; and attempts to understand the impact of innovation on growth and productivity are but a few examples. Patent counts have often been used in such research as one of the only observable, direct indicators of innovation occuring. However, it has long been recognized that the private and social value of innovation represented by a single patent varies widely (a contention supported by the results of this paper). This noise component of the patent count indicator has posed a significant obstacle to empirical analysis using patent data. Here, estimates derived from the model are used to construct a weighting scheme for patent counts based on the number of years that a patent is renewed. This removes 39 to 56% of the variance in value found in simple count data. Assuming that patent value is related to the value of the underlying innovation, such a weighting scheme can be used to improve the noise to signal ratio of patent counts as a

---

[1]Throughout, German and Germany refer to the Federal Republic of Germany before re-unification.

measure of innovation.

The results of the paper are also useful in understanding the role of patent systems and, in particular, for considering potential reforms. Patent rights are created by states largely to improve the incentives to innovate in the face of a market failure - the public good nature of information. Given this policy goal it is of interest to be able to quantify the magnitude of the benefits generated by the patent system. For instance, disagreement over the size of benefits was at the heart of the debate in the United States initiated by the pharmaceutical companies' claim that they no longer benefitted sufficiently from patent protection because the number of years between FDA new drug approval and the maximum patent term left them with too few years of monopoly sales. This debate resulted in the Pharmaceutical Extension Act of 1984 in the United States (and similar legislation in Japan). More recently, in discussions surrounding the Uruguay Round of GATT talks, multinational firms (and their national trade representatives) have claimed large losses due to weak protection when pressing developing countries to strengthen their intellectual property regimes. Independent evidence regarding the value of patent protection[2], such as that presented here, would improve the

---

[2]The value of protection measured here is the incremental returns obtained from an innovation over and above those obtainable using the method of appropriation which is the next best *given* that the inventor has applied for a patent. The value of the patent system is the incremental returns obtained with patent

empirical base upon which such issues are resolved.

Specific parameter estimates of the model also have interesting implications. For instance they indicate that patentees in some technologies are better able to defend their legal rights, that the development period (D as opposed to R) is essentially completed within 7 years, and that there is a fairly rapid rate of obsolescence of innovations. There is little other large sample empirical evidence on these points.

The estimation is based on two simple observations. First, a patentee must pay an increasing annual renewal fee to keep his patent in force. These range from 133 1975 DM for the 3rd year to 3,050 for the 20th and final year. The second observation is that

---

protection over and above those obtainable using the next best method available before application. There are two reasons why the *ex-post* and *ex-ante* best alternatives may differ. *Ex-post*, secrecy is no longer a viable option due to the disclosure requirement. On the other hand, the ability to appropriate returns via marketing/brandname power may be enhanced with time as a patentee monopolist. The importance of these two factors determines whether the value of protection as estimated here from renewal data is larger or smaller than the value of the patent <u>system</u> as a protective mechanism. (Patents may be valuable quite apart from their role enhancing appropriability because they allow information about innovative success to be publicized without constraint. This may make a positive contribution to the profits derived from an innovation insofar as publication improves employee morale, increases stock prices and improves the ability to obtain external finance.)

most patents (65-95%) are not renewed for the maximum number of years. Assuming maximizing behavior, the decision to renew in a given age depends on the returns to protection in that age plus (since non-renewal is permanent) the expected value of maintaining the option to continue in the future. A dynamic stochastic discrete choice model is developed and, in conjunction with renewal fee schedules, the implications for aggregate behavior are derived[3]. Variation in renewal fees both across age and across cohorts identifies the distribution of patent value. Because the model is analytically intractable, it is fit using a simulated minimum distance estimator.

The most significant departure from earlier modelling of patent renewal is in the recognition here that litigation may be important. Previous models of renewal behavior have not incorporated the fact that, while patent rights are granted by the state, they must be defended by the patentee. If a patentee will not defend his patent then others may infringe with impunity, and returns to protection are zero. Evidence presented here (section 2.4) and in Lanjouw (1992) suggests that the threat and occurrence of litigation does affect patentee decisions.

---

[3]The first use of renewal data to estimate the private value of patents was Pakes and Schankerman (1978) using a deterministic model of returns. A stochastic model was first formulated in Pakes (1986).

Estimation is done using newly collected data consisting of renewal information for a sample of West German patents from cohorts (application dates) 1953-80 and years 1955-88. The patents are from 3 nationality groups - Japan, United States, Western Europe - and 4 technology groups - computers, textiles, combustion engines and pharmaceuticals. Further details about the data are in Appendix I and a summary of the patent granting procedure, the various fees and timing of legal protection are in Appendix II. This paper presents the first patent value estimates disaggregated by type of technology and the results suggest that the profile of returns to patent protection varies substantially in this dimension.

2. The Patent Renewal Model

This section first presents the renewal decision rule for patentees. The stochastic element in this model is the positive or negative evolution of returns over time. In the most general form, the probability distribution of lifetime returns can be specified as vectors of returns for each age with probabilities attached. In order to use concepts with theoretical and empirical content, such as learning and depreciation, to inform both the construction of the model and the interpretation of the results, this vector probability distribution is cast in the form of a set of conditional probability distribution functions, where the pdf for each age is conditional on returns received in the previous age. A model of the evolution of returns over the life of a patent is described in subsections 2.2 and

2.3. Because patent rights must be defended by the patentee, the value of owning a patent depends on the ability to defend it. The threat of infringements are incorporated into the decision rule in the final subsection.

2.1 The Decision Rule

The value of renewing a patent in any age is equal to the returns to protection for that age plus the expected value of the option to continue renewing in the future minus the renewal fee, c_a. $\bar{r}_a$ is the minimum level of returns in age a which will lead to renewal. Thus, letting r_a^* be defined implicitly by the solution to:

$$V(a, r_a) = r_a + \beta\theta E[V(a+1)|r_a,\underline{c},\underline{\omega}] - c_a = 0, \qquad (1)$$

$$\bar{r}_a = r_a^* \text{ if } r_a^* > 0$$

and

$$\bar{r}_a = 0 \quad \text{if } r_a^* \leq 0$$

where $1-\theta$ is the probability of obsolescence, $1-\beta$ is the real discount rate (set to .05), $\underline{\omega}$ is a vector of parameters and $E[V(a+1)|.]$ is the expectation of the value of the patent in a+1 conditional on current returns and the vector of parameters, $\underline{\omega}$, assuming optimal future decisions. Where appropriate this is shortened to EV(a+1). Two points may be noted from the implicit definition of $\bar{r}_a$. First, because $E[V(a+1)|r_a,.]$ is continuous and non-decreasing in r_a, there is a unique solution. Second, $E[V(a+1)|r_a,.]$ is non-increasing and c_a is non-decreasing in age[4] which together imply that $\bar{r}_a$ is non-decreasing in age. (Proofs are in Lanjouw, 1992, Chapter 4)

2.2 The Evolution of Returns to Patent Protection - γ, σ, ϕ

At the time of renewal for age a, agents are assumed to know their returns for age a[5]. They are uncertain about the future evolution of their own returns. However, they know the probability distribution over all future events, conditional on their current

---

[4] There are, however, relatively high lumpsum costs due near the time of granting. This truncates the distribution of those granted such that the probability of renewal is higher in the ages just after granting than for patents of the same age not recently granted. To avoid 'lumpsum' effects in the sample hazard rates, the data is subsetted so that only those patents renewed a sufficient number of years past payment of the examination and publication costs are included in the calculation of sample hazard rates (where sufficient is based on preliminary estimations in a deterministic framework). 15% of the observations are lost in the process. However, computing model hazards conditional on both age and proximity to the payment of lumpsum fees would raise the number of computations at least five-fold. Further, the distance between granting and lumpsum fees is variable and unknown for individual patents and therefore would require intergrating out over another unknown distribution. Finally, subsetting has the added advantage of removing any distortions in post-grant hazard probabilities arising from litigation (see Lanjouw, 1992b).

[5] This is a reasonable assumption since renewal fees are not due until several months into the age in question and, moreover, returns are often the result of earlier contracts.

returns. This conditional probability density function is modelled as:

$$r_a = \max\{ z, \delta r_{a-1}\}$$

where $r_0 = 0$, $1-\delta$ is the depreciation rate and z is a draw from an exponential density:

$$q_a(z) = [\exp\{-((z/\sigma_a)+\gamma)\}] / \sigma_a.$$

$\gamma \geq 0$ and σ_a is modelled as $\phi^{a-1}\sigma$ with $\phi \leq 1$.

This specification for returns has the following characteristics. First, and of importance to deriving the decision rule, is that returns in any age are conditional only on returns in the previous age[6]. Second, with $\gamma > 0$ there is a mass point at δr_{a-1}. Returns at the point of application are 0 and there is therefore a positive probability of receiving returns of zero through any age a of $(1-\exp\{-\gamma\})^a$. This feature incorporates the often early date of applications and delays of varying lengths before commercial exploitation. The

---

[6] Returns in the previous age can be thought of as a sum of all of the revenue generating opportunities found up until that point. Then the probability $z > \delta r_{a-1}$ represents the probability of finding new markets in addition to those already found. With a limited number of potential customers, this declines as the number already found increases. One implication of this model of returns is that, for any $s > \delta k$, $\text{Prob}(r_a \geq s | r_{a-1} = k) = \text{Prob}(r_a \geq s | r_{a-1} = 0)$. This suggests that there is a fixed and finite number of opportunities, all of which may be found and begin to generate revenues in the first period, but which are more likely to be found only over time. A model where only a restricted number of opportunities are searched each age, would be additive with z distributed, say, exponentially, $y=\max\{0,z\}$ and $r_a = r_{a-1} + y$.

fact that $\sigma_a \leq \sigma_{a-1}$ captures the idea that patentees try those marketing opportunities that they expect to be most lucrative first and that in this regard they do at least as well as a random search ($\sigma_a = \sigma$). It also captures the idea that, as lead time passes, imitators may make inroads capturing some prospective customers.

The conditional distribution functions describing returns are common to all agents (in a technology group). *Ex-ante*, all patents are treated as covering innovations of the same quality. Individual heterogeneity evolves over time as the potential of some innovations is realized to a greater degree and/or earlier.[7]

---

[7] It would be interesting to introduce individual heterogeneity in *potential* with a distribution over σ. Then γ could continue to be interpreted as industry-linked marketing uncertainty, and the distribution of σ could be interpreted as ex-ante variation in the quality or potential of an individual innovation. Such a formulation would allow a patent generating no returns but with a high σ (low potential) to be dropped while a patent generating no returns but with a low σ to be renewed. Given sufficient variation in cost schedules, a distribution over σ would be identifiable. Differences in the pattern of non-renewals among patents for innovations not yet exploited and generating common returns of zero (ie, pharmaceuticals still under testing) might be informative about ex-ante variation in how valuable patent protection is expected to be. A further extension would allow patentees to be uncertain about the σ of their patents with learning over time as returns are realized. This change would be a formidable leap in the computational burden as each patent would have a different vector of $\bar{r}$'s based on its realized history of returns.

2.3 Decay in the Returns to Protection - δ, θ

Returns are subject to two distinct factors which may cause them to decline in age. The first is a constant annual rate of depreciation, $1-\delta$. The argument that returns depreciate in age is based on the fact that other innovations will arise, either directed or by chance, which can compete with the older one. Alternative methods of protection, such as branding, may be established. There may also be low levels of imitation not substantial enough to prosecute which diminish the returns to protection.

The second possible factor is obsolescence, which occurs with probability $1-\theta$. This is defined as a permanent fall to a level of returns such that the patent will not be renewed, even in the midst of an infringement suit. In essence, obsolescence is simply an extreme level of depreciation. In any age there will be some distribution of depreciation rates, δ_i, over a group of patents. δ in the model corresponds to the mean depreciation of those which do not become obsolete, while $1-\theta$ may be thought of as the percentage experiencing a very high level of depreciation. The assumption being made is that the distribution of depreciation rates is concentrated in a narrow range at the low end of the [0,1] support with, perhaps, a thin density at high levels, and so can be well summarized by the mean once the those at the high end are removed.[8]

2.4 Litigation - w, α_0, α_1

The possibility of challenges introduces a new element into the decision rule which operates as an additional cost to maintaining patent protection. Assuming common knowledge, if a patentee will not prosecute if all potential competitors infringe, then they will do so and returns to protection are zero. Therefore, the patentee will only renew if he is willing to defend the patent against infringers, which entails being willing to pay both legal fees and the patent office renewal fees due during an infringement suit. Most suits are completed within 3 years (US Department of Energy, 1982). $\bar{r}_a$ in a model incorporating the fact that a patentee must be willing to

---

[8] In a model where current returns are *not* known at the time of decision making and the rate of individual depreciation is uncertain, obsolescence would capture 'net mistakes'. For example, suppose that in any age, $\delta_i = \mu_\delta + \epsilon_i$ and ϵ_i is unknown when the renewal decision is made. Then the decision rule incorporates a joint distribution for (z,δ_i). (Note that it is not correct to just use μ_δ, the intuition being that decisions are based on probabilities of being above $\bar{r}_a$.) In age a+1, some agents will discover that they should not have renewed in a and will drop in a+1. On the other hand, some agents who should drop in a+1 will mistakenly renew. Obsolescence would appear if the distributions of ϵ and of r_a around $\bar{r}_a$ are such as to make the net mistakes attributable to uncertainty in δ positive. This interpretation of obsolescence is inconsistent with the informational assumptions of the model and it seems unlikely that net mistakes of this type could explain the levels of obsolescence suggested in Table 1. However, this type of effect may have some role in explaining part of θ.

defend against infringement is defined in the same form as above. Letting r_a^* be the unique r_a which solves the equation:

$$V(a, r_a) = \tag{2}$$

$$w\theta^2 r_a + w\beta\theta^2 E[VL(a+1)\,|\,r_a,.] - c_a - \beta\theta c_{a+1} - (\beta\theta)^2 c_{a+2} - (1-w)[LF] = 0,$$

then $\bar{r}_a = r_a^*$ if $r_a^* > 0$

and $\bar{r}_a = 0$ if $r_a^* \leq 0$

$E[VL(a+1)\,|\,.]$ is the expected value of the future if two years remain before resolution of a suit, LF represents legal fees, and w is the probability that the patentee successfully defends his patent[9].

An important feature of challenges is that returns during the period of litigation are contingent on winning the suit (ie, in the

---

[9] This formulation of equilibrium is the simplest way to introduce litigation into the model and is dependent on the common knowledge assumption. This assumption seems reasonable here since all agents are likely to have similar knowledge about the market and, due to the disclosure requirement, similar knowledge of the innovation. One could posit other models of equilibrium. If there is uncertainty in the minds of potential challengers as to whether a patentee will defend, then the patentee has an incentive to renew in order to dissuade them. Even with common knowledge, if a patentee expects to have a series of patents then he might find it advantageous to randomize his defense in order to establish a reputation as an aggressive defender of his property rights and lower the probability of challenge, say P^c, below one. $\hat{w}$ estimated here equals $\{(1-P^c) + w^*P^c\}$ which is greater than the true probability of winning, w, under either of the alternative specifications.

form of damages). If the patentee abandons the suit then clearly he does not win. The fact that returns during a suit are contingent on renewal throughout the suit has the following important implication for determining $\bar{r}$'s. Decisions in a+1 and a+2 *during a suit* now depend on earlier returns: $\bar{r}_{a+1}'(r_a)$ and $\bar{r}_{a+2}'(r_a, r_{a+1})$ where ' denotes $\bar{r}$ during a suit. That is, a patent may be renewed in a+2 in order to secure returns from age a. In particular, while in the absence of litigation obsolescence always leads to non-renewal, during a suit a patent experiencing obsolescence in a+1 or a+2 may continue to be renewed in order to obtain the earlier returns. Taking into account the dependence of $\bar{r}_{a+1}'$ and $\bar{r}_{a+2}'$ on previous returns and ensuring dynamic consistency when solving for $\bar{r}_a$ is theoretically straightforward[10]. However, as a practical matter, doing so adds very substantially to the computer time necessary to calculate the $\bar{r}$'s (which must be done for every cohort at every iteration). The renewal rule (equation 2) incorporates two assumptions made in order to avoid the computational burden. First, the option is sufficiently important that obsolescence arriving either in a+1 or a+2 will lead to non-renewal. Note that where this assumption is not correct, $\bar{r}$ is biased upward and the only patents affected are those with returns falling between $\bar{r}_a(1)$ and $\bar{r}_a(2)$. Second it is assumed that, barring obsolescence, an infringement will be

---

[10] The procedure is to specify r_a and r_{a+1} and step search for $\bar{r}_{a+2}'$. Then repeat in a step search for $\bar{r}_{a+1}'$ given r_a and $\bar{r}_{a+2}'$. Finally calculate the value function $V(r_a, a)$ given $\bar{r}_{a+1}'$ and $\bar{r}_{a+2}'$and repeat in a step search to find $\bar{r}_a$. The nested nature of these step searches is the factor which magnifies the computational time.

prosecuted through the future 2 years. The bulk legal fees have to be paid even if a suit is abandon in midstream so both c_a (and c_{a+1}) and legal fees are sunk costs. It seems plausible to assume that in most cases these sunk costs are large enough relative to the distance between the worst news about the value of continuing and the expectation at age a that the patentee would continue. Together these two assumptions imply that, during an infringement suit, $0<\bar{r}_{a+1}' \leq \delta r_a$ and $0<\bar{r}_{a+2}' \leq \delta^2 r_a$.

Attorneys' fees and court costs are borne by the losing side. Their levels are set by statute and are linked to the damage inflicted on the patentee - as measured by lost profits. Annual damages are, by definition, exactly equal to the returns derived from protection from competition. The size of total legal fees depends on the value of the patent and the duration of the process. Information from several sources (i.e., Berkenfeld, 1967; Bohlig, 1987; Korner, 1984) indicates that the total legal fees increase approximately linearly in damages:

$$\text{Total legal fees} = \alpha_0 + \alpha_1 r_a$$

with OLS estimates of $\hat{\alpha}_0 = 12{,}612$ and $\hat{\alpha}_1 = .20$ (1975 DM, estimated standard errors of 5001 and .02 respectively, $R^2 = .89$). These estimates are used throughout.

Clearly anyone can spend more than the statutory fees on legal services. One might model w as an endogenous function of the money spent by the patentee to prosecute the suit, LF^p, and that spent by the challenger, LF^c. For example, suppose that both sides have access to equivalent legal services (in the sense that $\partial w/\partial LF^p = - \partial w/\partial LF^c$) and, for simplicity, that they would face the same benefit from winning the suit, B. One can model this situation as a simple one-stage game with continuous strategies where, for each player, the strategy variable is how much to spend on legal services over and above the statutory legal fees. Assuming convexity in the function $w(LF^p,LF^c)$, the equilibrium will be symmetric at a level of LF such that the increased probability of winning from an incremental increase in LF does not warrant the expense. As $B \rightarrow \infty$ so too does the equilibrium level of legal costs.[11] This could present a problem in specifying the legal costs that a patentee expects to incur if challenged. However, the problem does not arise here because all that is of importance in the decision rule is the equilibrium legal fees for those on the <u>margin</u> of renewing. It suffices to assume that $w(LF^p,LF^c)$ is sufficiently convex that patents with low returns will not bring forth expensive legal battles and that patents with high returns do not become marginal.[12] As a result,

---

[11]Ashenfelter (1990) has an interesting discussion and empirical examples of prisoners' dilemma incentives arising with respect to lawyers in several arbitration settings.

[12]Another reason to expect minimum legal costs is that there may be some coordination among the parties to avoid mutually damaging expense. That this is common is suggested by a study of German patentees conducted by the US Department of Energy (1982). They

in the derivation of the renewal decision rule, legal fees are set using the statutory fee schedule.

Legal decisions regarding the validity and scope of a patent are made on a technical rather than an economic basis, so there is no reason to expect that the probability of winning a challenge is related in any way to the value of the patent or to expect that it changes over age. If patentees have differing w, their expectations of winning, then their behavior will reflect this difference in the same way as known individual differences in δ. However, given that all patents undergo an extensive examination before granting it does not seem unreasonable to assume that w is the same across patents within the same technology group.

The thoroughness of the examination procedure suggests also that w should be rather high. Two other facts also point in this direction. First, except where a challenge is by another inventor claiming priority, the benefits to a challenger of winning a suit are less than those to the patentee because the challenger can get access but not exclusive use of the innovation. Hence, considering the simple game in legal fees described above, with B not equal equilibria are likely to be asymmetric with the patentee investing

---

found that large corporations rarely litigate and that overall more than a third of suits are settled out of court. Of course settlements entail legal costs, but the fact that settlements are regularly made suggests that they are less costly.

somewhat more in the battle with a higher resulting probability of winning. A second indication is the very low numbers of patents which are revoked. Well over 50% of all infringement suits involve revocation proceedings. Of 101 revocation suits initiated in 1980 (.07% of those in force), under a third led to revocation (US Department of Energy, 1982). A larger study of all (194) revocation defenses initiated in 1972-74, found that just 17% led to revocation of the patent (Stauder, 1983). Since revocation suits are initiated by the challenger, insofar as w does vary across patents one would expect them to occur against patents which are more vulnerable. This suggests that overall, patentees have a higher probability of winning infringement suits than that indicated by observed success rates. (However, this must be weighed against the fact that patents with low w may not be defended and hence will not contribute to observed success statistics).

It is not necessary to distinguish between those patents which are revoked and those which are dropped according to the specified renewal rule because only .022% of the patents in the data were revoked by 1988[13]. This minute number is treated as zero. The

[13]The fact that the number revoked is so small does not mean the there are few infringements. A large number of challenges are settled out of court, usually with a licensing agreement. From the point of view of the decision rule, as long as the settlement terms correctly reflect the patentee's probability of winning (as evidenced by those which do go to court) and thus give him a stream of royalty payments equal to his expected stream of benefits from pursuing the suit, the two events are equivalent in the eyes of a risk neutral

threat of infringement has an indirect effect on hazard rates via the decision rule but does not have a significant direct effect.

It is assumed in the model that the patentee must be prepared to be challenged throughout its life, whether or not it has already been challenged and successfully defended. This is because there are likely to be new sets of potential infringers or types of infringements as new ways to commercially exploit the innovation or new competing innovations arise over time. In his study of infringement suits brought 1972-74, Stauder (1983) found that about 15% were challenged twice, 3% three times and several had four successive challenges. The point of this assumption is that it makes the same decision rule apply to all patents - whether they have been challenged before or not.

A final question that must be considered before the model can be estimated is what expectations patentees hold regarding future renewal fees and patent term length. The 1977 increases in the annual fee schedule and the maximum term length allow the data to suggest an answer. The higher renewal fees came into force in 1977 and hence applied to renewals for the following year, 1978. The renewal fees were increased substantially, by 48 to as much as 81% depending on the age. At the same time, the maximum term

---

patentee (a risk averse patentee would be willing to accept lower royalties in return for the added certainty).

increased from 18 to 20 years. In a deterministic model with decreasing returns, where the future does not enter renewal decisions and $\bar{r}$'s equal c's, this change would unambiguously cause hazards for all ages to increase in the year 1978. As shown in Table 1, while for some ages there was a higher hazard in 1978 compared to 1977, the changes were by-and-large small and for several ages the hazard actually fell. This observation suggests strongly that the future does indeed enter the decision to renew thus justifying the stochastic framework used here. Moreover it indicates that cost changes are anticipated, leading to the assumption that patentees have correct expectations about the renewal fee schedule that they will face for future ages[14].

When the future matters and changes in renewal fees are anticipated, the hazards at the time of the change can be either relatively high or low, the reason being that anticipated cost changes also effect $\bar{r}$'s for the previous year. The effect on hazards in the year of change depends on the relative changes in the two vectors of $\bar{r}$'s and the distribution of returns. A striking feature of the data is that, in 1979, the hazards fell dramatically (Table 1). In particular, it is the late age where hazards which are relatively low, a fact which indicates that the reason may be the change in term length. When

---

[14] Annual fees include those due to the German patent office as well as an administration cost of 85 1975 DM. Patent administration is a marketed service and therefore can be priced. Information was obtained from Ms. Angela Wycherly of Computer Patent Annuities, Jersey, UK.

Table 1

Impact of the 1977 Fee & Term Length Changes1

Age	Hazard78/77	Hazard79/77	Hazard80/77
6	1.00 (.31)	.81 (.26)	.56 (.21)
7	1.71 (.44)	.90 (.27)	1.16 (.31)
8	1.03 (.23)	1.30 (.27)	.96 (.22)
9	.86 (.20)	.90 (.20)	1.12 (.24)
10	.57 (.14)	.81 (.16)	.69 (.14)
11	.94 (.17)	.73 (.13)	.74 (.13)
12	1.05 (.19)	.96 (.17)	.63 (.13)
13	1.05 (.16)	.46 (.10)	.87 (.14)
14	1.09 (.20)	.26 (.08)	.86 (.16)
15	1.42 (.25)	.22 (.09)	.62 (.13)
16	1.70 (.33)	.24 (.11)	.85 (.20)
17	1.13 (.26)	.25 (.13)	.81 (.22)
18	.94 (.16)	1.11 (.18)	.42 (.12)

Notes:
1) All hazard rates are averages over nationality/technology weighted by sample sizes for the given age.
2) Standard errors estimated using a Taylor expansion are in parentheses.
3) The same trends were found in the technology groups taken separately.

the future matters, an extension of the statutory patent life increases the expected future value of renewing and hence lowers $\bar{r}$'s (except, perhaps, in ages 17 and 18, see below). Furthermore, the impact is greatest at those ages near the end because, having survived until that point, patents are more certain to continue through 18 and benefit from the extension. As in the case of the cost changes, changes in term which are anticipated can lead to either relatively high or low hazards. An unanticipated term change, on the other hand, would lead to unambiguously lower relative hazards such as are observed in the data. Because the fall in hazards is so distinct, it is assumed that the term change was not anticipated, and further, since the impact is seen in 1979, it is assumed that it takes a year for such information to be incorporated into patentee decision making.

The fact that the observed hazard for age 18 is substantially higher in 1979 is evidence in support of the contention that the threat and occurrence of infringements influences patentee decision making. In the absence of the threat of infringement, expected future value can only increase in the maximum term length (patentees can always decide not to avail themselves of the extra ages). However, with the threat of infringement, renewal in age 18 requires a willingness to continue for ages 19 and 20, the cost of which increases from 0 to a relatively high level with the increase in maximum term. Similarly, renewal in age 17 requires a willingness to pay the new renewal fee for age 19. With certain sets of costs and parameter values the extension of term length may no longer be a positive event. In such a case, $\bar{r}_{17}$ and/or $\bar{r}_{18}$ increase in the maximum term. The fact that the observed hazard for age 18 is much higher in 1979 relative to other years suggests that the threat of litigation does have a discernable negative effect on expected value and that protection in ages 19 and 20 does not generate positive net returns for those on the margin of renewing for ages 17 and 18.

3. Estimation

A recursive system of equations for finding the minimum returns required for renewal in each age, the $\bar{r}$'s, is given in Appendix III. A recently developed approach to the estimation of models which are analytically intractable is to simulate the model at a give set of parameters rather than solving for an exact solution to the objective function. The estimator used in the following estimations of the model is a weighted simulated minimum distance estimator. The following subsection discusses briefly some properties of the estimator. The second discusses the estimations and convergence.

3.1 Properties of the Estimator

The model is estimated using a simulated minimum distance estimator, $\underline{\hat{\omega}}_N$, of the true parameter vector, $\underline{\omega}_0$, where $\underline{\hat{\omega}}_N$ is chosen so as to minimize the norm of the distance between the vectors of sample hazard proportions, $\underline{h}_N$, and those which are simulated.[15]

---

[15] Hazard rates are the appropriate form in which to use data on granted patents in situations where granting occurs over an extended number of ages as in Germany. The reason is that renewal proportions and mortality rates will both have a selection bias introduced by the fact that most patents which are dropped in the

The subscript N denotes the sample size. $\underline{\hat{\omega}}_N$ is a vector of parameters which are 'sufficiently' close to the global minimum of the objective function:

$$\| \underline{G}_N(\underline{\omega}) \| = \| \underline{h}_N - \underline{\eta}_N(\underline{\omega}) \| \tag{3}$$

where $\underline{\eta}_N(\underline{\omega})$ is a simulation estimate of the aggregate hazard rates implied by the parameter vector $\underline{\omega}$. The hazard vectors are of dimension m, equal to the number of cohort/age cells. More formally, it is assumed that the optimization procedure minimizes $\| \underline{G}_N(\underline{\omega}) \|$ such that:

(i) $\| \underline{G}_N(\underline{\hat{\omega}}_N) \| \leq \inf_{\Omega} \| \underline{G}_N(\underline{\omega}) \| + o_p(N^{-1/2})$.

Identification, uniformity and continuity conditions required to ensure the consistency and asymptotic normality of this estimator may be found in McFadden (1989) and Pakes and Pollard (1989). Proofs that some of these conditions hold for a similar patent renewal model are found in Pakes and Pollard, and Pakes (1986). Whether these are appropriate for the model presented here is

---

early ages never appear in the granted population due to the extended granting period. The result is that renewal proportions are biased upwards and mortality rates are biased downwards. This is true even though patents granted at different ages have the same mortality probabilities when viewed from age 1. (Chi-square tests could not reject a null hypothesis of equal hazard probabilities for patents of different grant ages in these data) Hazard rates, because they are conditional on being alive in the previous age, are unbiased.

considered for each in turn and in most cases proofs are based on those found in the above cited papers.

It is assumed that $\underline{\omega}_0$ is a unique point in Ω such that $\underline{G}(\underline{\omega}_0) = \underline{0}$. In addition to condition (i) above, consistency of the estimator requires two additional conditions which ensure that $\underline{G}_N(\underline{\omega}_0)$ is eventually close to zero and that $\|\underline{G}_N(\underline{\omega})\|$ is small only near $\underline{\omega}_0$.

(ii) $\underline{G}_N(\underline{\omega}_0) = o_p(1)$.

One can see that this condition is fulfilled by noting, as do Pakes and Pollard, that

$$\|\underline{G}_N(\underline{\omega})\| \leq \|\underline{h}_N - \underline{\eta}(\underline{\omega})\| + \|\underline{\eta}(\underline{\omega}) - \underline{\eta}_N(\underline{\omega})\|.$$

Since the samples are random, the observations are independent[16] as are the simulation draws. Therefore the law of large numbers ensures that $\|\underline{G}_N(\underline{\omega}_0)\| = o_p(1)$.

(iii) $\sup_{\|\underline{\omega}-\underline{\omega}_0\| > \delta} \|\underline{G}_N(\underline{\omega})\|^{-1} = O_p(1)$ for each $\delta > 0$.

---

[16] Several patents may be taken out to cover different aspects of a single innovation. For this reason, such patents may be dropped as a unit. Such behavior does not affect the sampling distribution if the group is treated as a unit. However, since links between patents are not available in the data, each patent is treated as independent of the others. The variance derived in Appendix IV for the sample hazards assuming independence understates the variance where patents are linked.

This condition will be satisfied given sufficient variation in the renewal fee schedule over age and time.

Once it is shown that $\underline{\hat{\omega}}_N$ is a consistent estimator of $\underline{\omega}_0$, Pakes and Pollard also show that, the following conditions (iv) - (vi) ensure that if $\underline{G}_N$ converges in distribution (denoted by $\rightarrow$) to an m-variate normal,

$$\sqrt{N}\ \underline{G}_N(\underline{\omega}_0) \rightarrow N_m(\underline{0},V)$$

then $$\sqrt{N}\ (\underline{\hat{\omega}}_N - \underline{\omega}_0) \rightarrow N_k(\underline{0},\ (\Gamma'\Gamma)^{-1}\Gamma'V\Gamma(\Gamma'\Gamma)^{-1}), \tag{4}$$

where Γ is an m by k derivative matrix of $\underline{G}(\underline{\omega}_0)$ of full rank. The distribution of $\|\underline{G}_N(.)\|$ is derived in Appendix IV.

(iv) $\underline{\omega}_0$ is an interior point of Ω, where Ω is the k-dimension space over which the parameters are optimized.

This condition may or may not be satisfied. All of the parameters are bounded at least from below by zero. There are narrower bounds on some of the parameters due to non-parametric results (θ - see bounds in the following section) or by definition (w, δ, ϕ - upper bound of one).

(v) $\underline{G}(.)$ is differentiable at $\underline{\omega}_0$ with a derivative matrix of full rank. (For a proof see Lanjouw, 1992a, Chapter 4.)

(vi) for every sequence $\{\delta\}$ of positive numbers that converges to zero:

$$\sup_{\|\underline{\omega}-\underline{\omega}_0\| < \delta_N} \frac{\| G_N(\omega) - G(\omega) - G_N(\omega_0) \|}{N^{-1/2} + \|\underline{G}_N(\underline{\omega})\| + \|\underline{G}(\underline{\omega})\|} = o_p(1)$$

Pakes (1986) proof that condition (vi) is fulfilled for his patent renewal model uses the form of the conditional distributions of returns over age and the fact that the renewal decision hinges on $\bar{r}$'s. It does not depend on what the $\bar{r}$'s look like and therefore the same proof applies here. One important note: the proof of the equicontinuity condition (vi) requires that the difference between the simulated and true model, the simulation bias, must decrease as $\underline{\omega} \rightarrow \underline{\omega}_0$. This can only hold if the same simulation draws are used throughout the optimization procedure.

As is true for other estimators, the efficiency of the simulated minimum distance estimator can be improved with an appropriate matrix of weights. The weighting matrix used here is

$$A_N(\underline{\omega}) = \text{diag}[\sqrt{[\underline{n}/N]}] \tag{5}$$

i.e., weighting by sample sizes. $\sqrt{[\underline{n}/N]}$ is the sample analogue of $\underline{\lambda}^1$ (see Appendix IV).

3.2 Estimations and Convergence

For each industry a grid search was conducted with a total of 216 combinations. These were studied for patterns and the best set of parameters from the grid search were used for estimating the function evaluation error, $\hat{\epsilon}_a$, and good finite difference intervals for the gradient following the procedure described in Appendix VI. Results from the grid searches were also used to inform the choice of starting values in the optimization routine.

Bounds were placed on each parameter to prevent the calculation of the function at unreasonable or undefined parameters. γ and σ were given a lower bound of zero. w, δ and ϕ were bounded between 0 and 1. θ was bounded between .90 and .95 (computers and pharmaceuticals) or .88 and .95 (textiles, engines). The lower bound reflects the fact that the obsolescence rate (1-θ) cannot be higher than the hazard rate and was checked against $\underline{h}_N$, the sample hazard proportions.

With each technology group a typical optimization run consisted of 5 or 6 iterations. The procedure ended in all cases with a zero finite gradient estimate. Perturbations of the parameter vector in both directions increased the function value. The estimation procedure was well behaved in the sense of moving quickly and directly to a local minimum. However, there are clearly many local minima. Given this fact, the procedure was restarted numerous times at various starting values near to the most promising areas as indicated first by the grid search and later by the results of further runs. The final estimates are those which yielded the lowest calculated value for the objective function.

4. Results of the Estimation and Implications for Patent Value

The final parameter estimates for each technology group estimated over the full range of cohorts and years are presented in Table 2, Part I. As expected, the estimated probability that a patentee will win an infringement suit, $\hat{w}$, is fairly high for all types of patents, particularly in pharmaceuticals. There is no observable counterpart to w, the probability of winning for all (not just challenged) patents. However, for comparison, recall that the two studies referred to in subsection 2.4 found aggregate patentee success rates in revocation proceedings of 69 and 83% - high, but still below the levels of around 90% suggested by the estimates of w here. Observed success rates *higher* than estimated w's would be difficult to reconcile. Observed success rates somewhat lower are reasonable once one allows that true w may differ across patents. While all patentees must be willing to defend, few are actually challenged. It seems likely that if some patents in a give technology group are more likely to be in danger of revocation then it is exactly those which are most likely to challenged - yielding an average probability of being successfully defended which is lower than that which would apply to the group as a whole.

A high w does *not* imply that the threat of infringement has no bearing on the renewal decision. A w of one does cause legal fees to fall out of the renewal rule - if the patentee is certain of winning he knows that he will not have to bear the legal expenses.

Table 2

Parameter Estimates[1]

I. By Technology Group - All Cohorts

	Computers	Textiles	Engines	Pharmaceuticals
w	.909	.931	.860	.989
	(.00151)	(.00194)	(.00489)	(.00228)
ϕ	.477	.574	.598	.306
	(.01330)	(.00964)	(.01685)	(.28260)
δ	.945	.959	.935	.942
	(.00057)	(.00055)	(.00130)	(.00116)
θ	.930	.923	.880	.900
	(.00119)	(.00321)	(.00570)	(.01047)
γ	.266	.530	.070	.000
	(.00370)	(.00534)	(.00834)	(.01267)
σ	4941.8	4086.8	9534.4	13111.2
	(365.92)	(690.99)	(624.94)	(1018.1)
size of:				
sample	5262	3526	3958	3029
simulation	5292	7052	7916	6058
MSE[2]	.00639	.00499	.00496	.00420
Var(h)	.01797	.02351	.02406	.06106

II. Two Period Estimates for Pharmaceuticals

	Period 1 (Cohorts 1953-66)	Period 2 (Cohorts 1967-80)
I. Parameter Estimates		
w	.900	.990
	(.00527)	(.00134)
ϕ	.400	.383
	(.05461)	(.01988)
δ	.975	.937
	(.00268)	(.00070)
θ	.913	.924
	(.01639)	(.00248)
γ	.000	.598
	(.00800)	(.00407)
σ	12114.0	8072.0
	(951.95)	(559.30)
size of:		
sample	1198	1831
simulation	4792	5493
MSE	.00068	.00112
Var(h)	.02185	.11530

Notes:

1) Estimated standard errors are in parentheses.

2) MSE is calculated as the weighted sum of squared residuals divided by 356, the number of cohort/age cells in the data. In Part II, the numbers of cohort/age cells are 210 and 146 for period 1 and 2 respectively. Var(h) is the variance in sample hazard rates.

However, it does not remove the (potentially) negative fact that the prosecution of an infringement requires payment of the renewal fees over the period of litigation. In fact, for the pharmaceutical group where $\hat{w}$ is close to one, the expected future value at each age for patents on the margin of renewing is negative for most ages, even though expected legal fees are negligible (i.e., $\bar{r}_a > c_a$). Without the threat of infringement, the expected future values would all be zero or positive. The threat of litigation effects the structure of the renewal rule and is not removed by setting w to one.

δ and θ describe the decay of patent value over time. Annual depreciation in returns is fairly slow - ranging from 4.1% for textiles to 6.5% for engines. However, for all patents there is a significant probability each year of obsolescence. Considering obsolescence in isolation, the probability of surviving through age 10 is only 32% for engine patents and at most 52%, in computers. Together these two parameter estimates imply quite a rapid decline in value over time. There is little other empirical evidence on this point. In one of the few studies, Mansfield, *et al.* (1981), in an in-depth look at 48 product patents, found that 60% of the patented successful innovations had been imitated (or rather invented around - legally) within 4 years of introduction.

γ, ϕ and σ together determine the path of learning. In particular, a high γ increases the probability of returns staying zero for some time. A low σ implies that opportunities are limited with a relatively small probability of a given patent being very valuable. A small ϕ means that potential learning opportunities are explored or lost quickly. Table 3 indicates what the parameter estimates from Table 2 imply about the speed of learning. The most striking finding is that all learning (net of depreciation) is over by age 6 or 7 in all technology groups[17]. The high value of $\hat{\gamma}$ in textiles is reflected in the fact that over 40% have zero returns in the first age. The engine group exhibits the largest gains in the level of returns over age - a reflection of a relatively high $\hat{\sigma}$ which decays at a relatively slow rate, $\hat{\phi}$ = .598.

The speed of learning about the value of an innovation is important in interpreting the significance of changes in the number of patent applications or grants. The decline in total patents and patents per R&D dollar over an extended period in the US, for instance, has raised concerns about diminishing returns in innovation activities. As pointed out by Griliches (1990), if patentees know the value of their patents at the time that they make application or renewal decisions, then a shift down in returns or a shift up in the cost schedule could lead to fewer applications and grants with little change in the average value since it would be the marginal ones which no longer proceed. On the other hand, if there is considerable uncertainty in the minds of patentees, then their decisions are not marginal but rather are based on expectations. If

[17] Pakes (1986) reports a similar result for aggregate German patents, with learning essentially over by age 5.

Table 3

Time-frame of Development and Commercialization[1]

I. Percentage of Patents with Increase in Returns in Each Age[2]

Age	Computers	Textiles	Engines	Pharmaceuticals
1	77.4%	58.9%	93.5%	99.2%
2	37.3	37.2	40.3	25.0
3	15.4	21.7	17.4	4.5
4	2.3	11.0	7.6	0.4
5	0.1	2.2	2.6	0.0
6	0.0	0.3	0.9	0.0
7	0.0	0.0	0.2	0.0
8	0.0	0.0	0.0	0.0

II. Percentage Increase in Returns over the Previous Age

Age	Computers	Textiles	Engines	Pharmaceuticals
2	869.1%	881.6%	2311.1%	1004.1%
3	168.4	286.2	578.1	53.5
4	10.8	91.2	154.7	1.4
5	0.4	9.6	28.9	0.1
6	0.0	0.6	6.8	0.0
7	0.0	0.0	0.6	0.0
8	0.0	0.0	0.0	0.0

Notes:
1) Calculations use 15000 simulation draws.
2) No increases in returns were found for ages greater than 7.

this is the case, to observe a fall in numbers suggests that there are fewer innovations being produced. The model here assumes that at the point of *application* patentees know only the probability density function over their possible future returns. Thus changes in patent applications fall under the second scenario and falls in their numbers would represent fewer total innovations. On the other hand, since learning is over by age 6 or 7, patentees (since 1967) can wait until the uncertainty is resolved before deciding whether to continue to granting. Thus, with respect to patents *granted*, falls in numbers fit the first scenario.

A surprising feature of the parameter estimates in Table 2 is the implication that pharmaceuticals have the most rapid development and commercialization, exactly the opposite of the result expected. Less than one percent of patents fail to achieve any returns in the first age. This is very difficult to believe and is very likely to have been imposed by the feature of the model that all patents are the same *ex-ante*. Suppose that the true data generating process is characterized by both a high level of returns and a delay in their commencement. The high expected values enter the decision rule for patentees making them willing to renew even if they are receiving no returns in order to maintain an option on the future. That is, $\bar{r}$'s are zero for a large number of ages and all patents are renewed except for those which become obsolete. If the delay in the arrival of returns is captured in the model by a large $\hat{\gamma}$, then a large number of patents remain with zero returns for many ages. Because of the imposed *ex-ante* homogeneity all patents receiving the same returns share expectations about the future, and therefore the same decision rule $\bar{r}$'s. As a result, in the first age that $\bar{r}$ moves above zero, all of those with zero returns drop together, giving an enormous simulated hazard for that age. There is some indication of this combination, high value with a delay, causing mass drops in the data at later ages. Looking in Figure 1.d at a weighted average of the sample hazard proportions, one can see that there is a jump in the hazards around age 10-11. Nevertheless, because the model does not allow for *ex-ante* heterogeneity, it can only generate a very extreme version of this mass drop out phenomenon and cannot successfully accommodate high returns which only begin at a late age. The model parameter estimates imply high returns, but avoids a massive drop by ensuring that essentially all patents have strictly positive value from the start.

The pharmaceutical group was explored further by allowing for differences over time. In particular, there are two reasons for believing that the gestation period for pharmaceutical patents was longer in the latter part of the period than in the former part. First, there has been a linear increase in the number of years required to obtain approval to market new pharmaceutical products. Second, there was an abrupt change in the population[18] of pharmaceutical

[18] Not only in type but in numbers. There were approximately 80 pharmaceutical patents granted per cohort before 1967, a number which jumped in 1967 to 150 and increased thereafter.

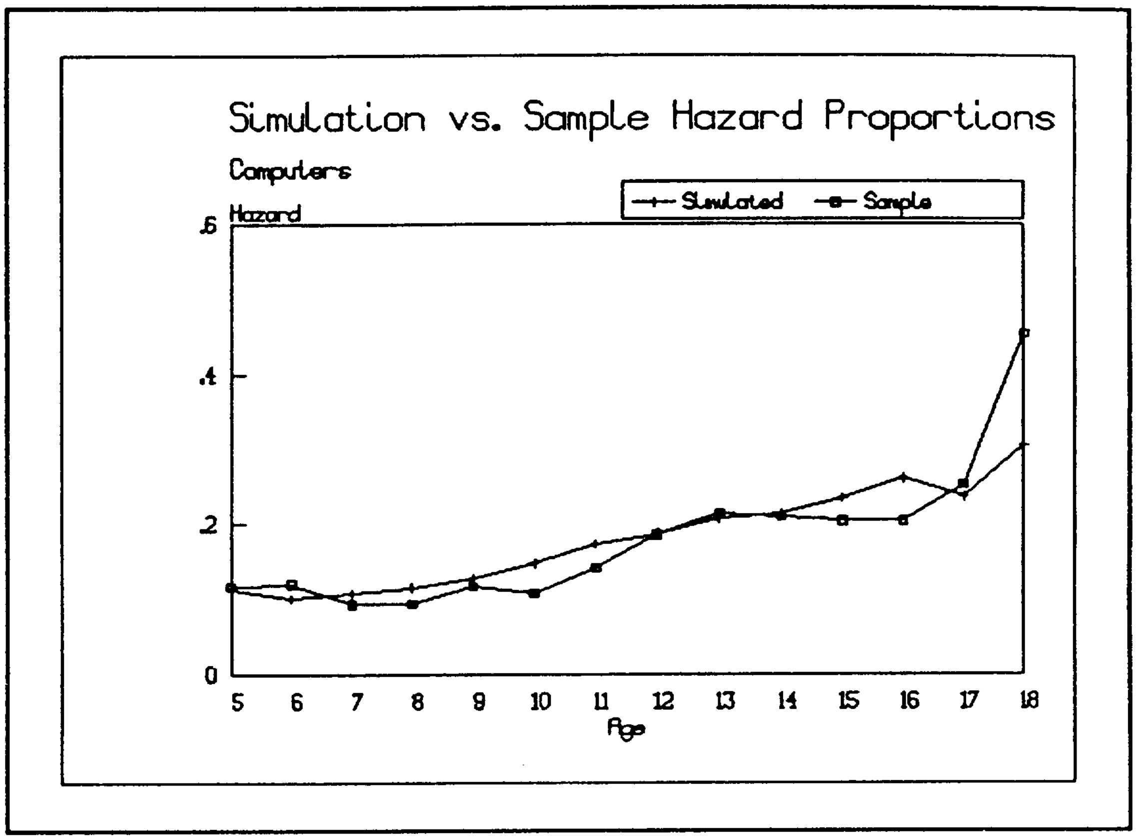


Figure 1.a

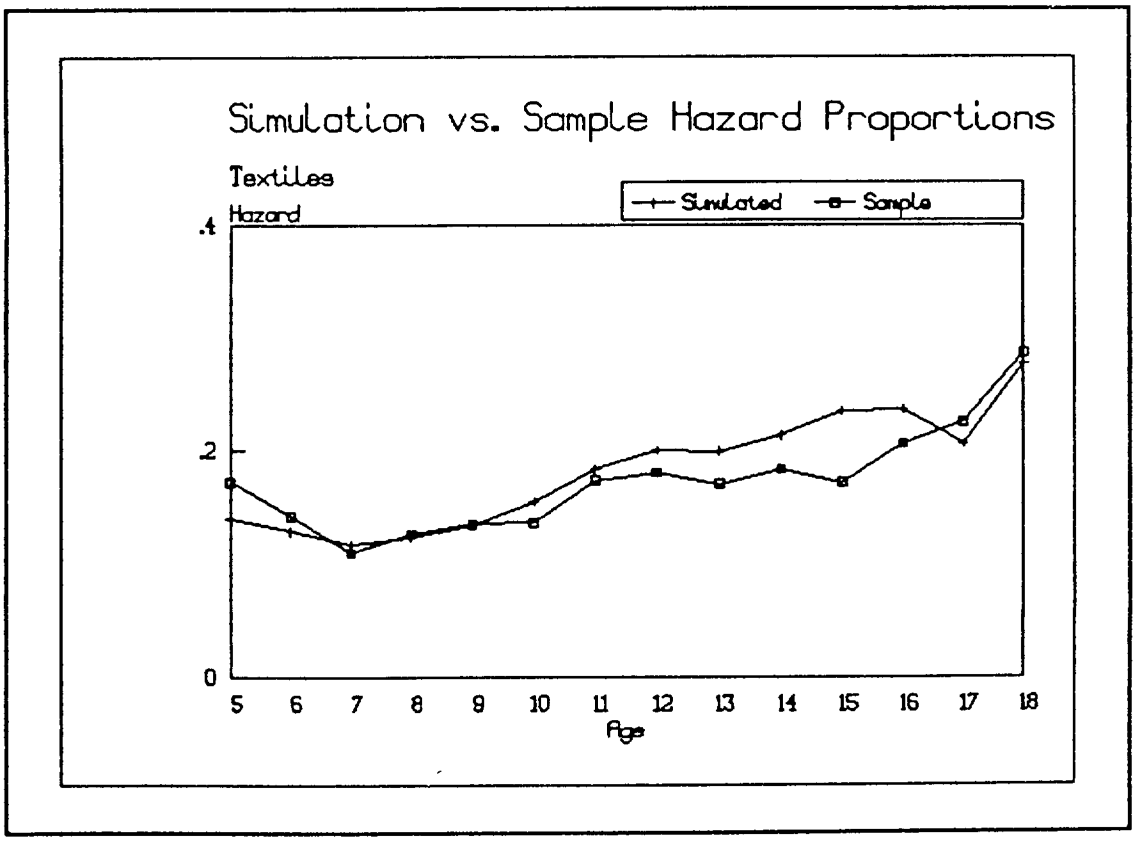


Figure 1.b

Figure 1.c

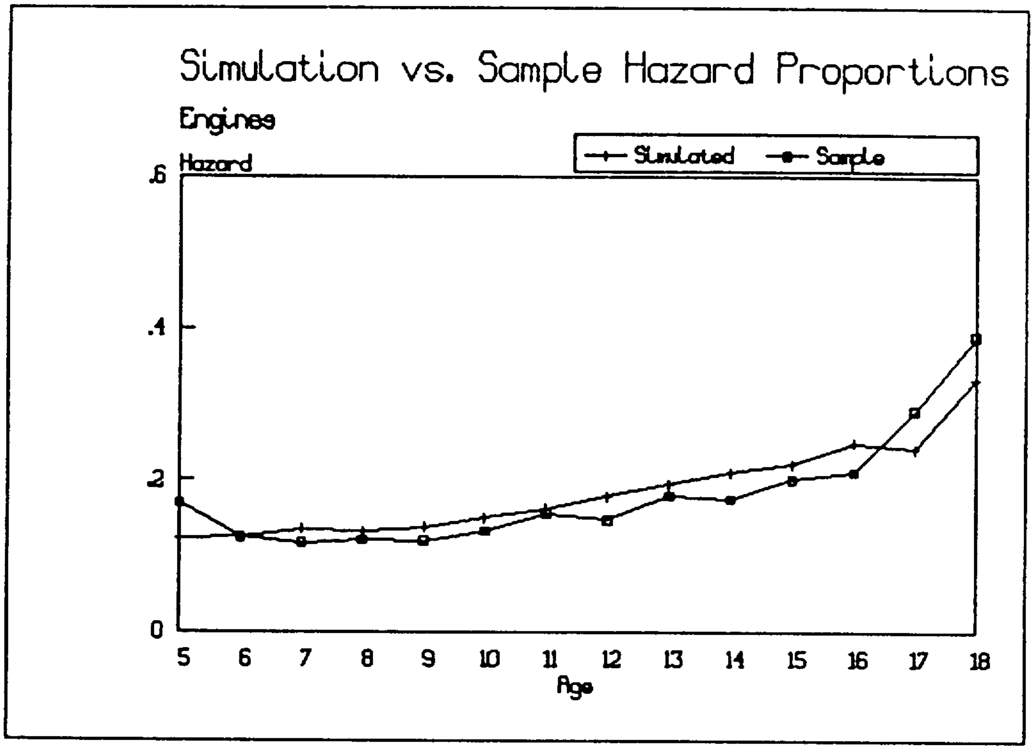
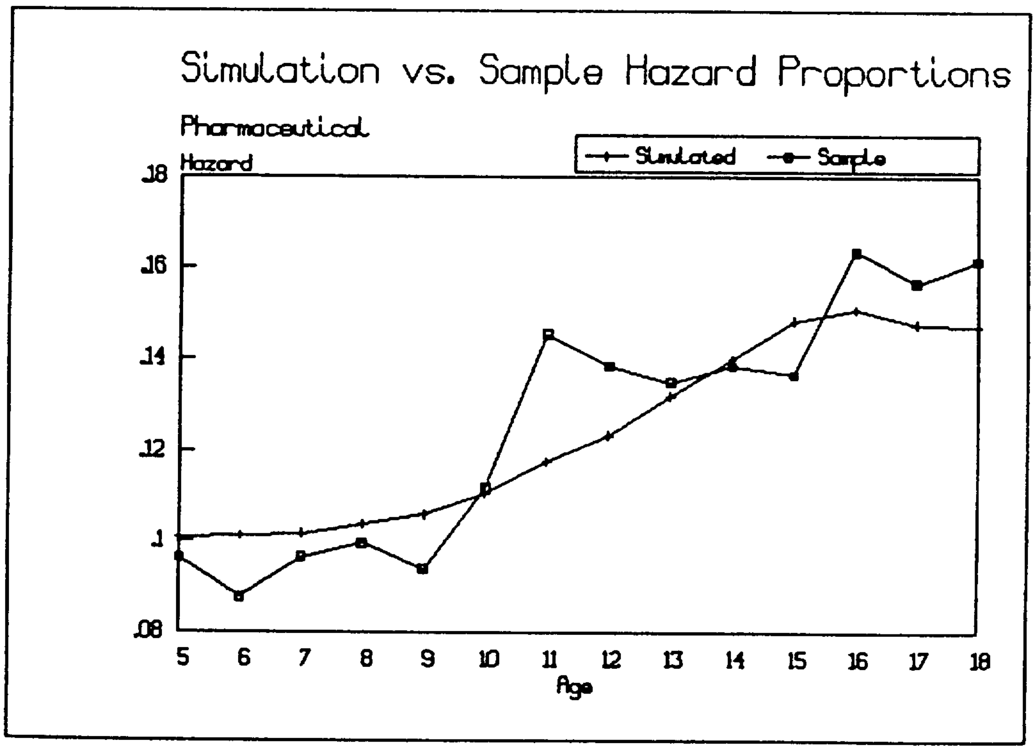


Figure 1.d



patents in 1967 when it became possible to patent <u>products</u> as well as processes. Processes, being internal to the firm, do not require the same marketing time. More importantly, it is only products which are required to undergo the German Health Agency testing. It seems quite plausible that the characteristics of patents from cohorts before 1967 differ markedly from those after. For this reason, the model was estimated for the two periods taken separately. Results are in Table 2, Part II. The two period estimates are almost all estimated with considerably more precision than the aggregate estimates, this despite the fact that both the sample and simulation sizes are smaller. This suggests that the two period break should be maintained and most of the analysis which follows does so. The estimates indicate that the probability of winning infringement suits increased substantially in the second period. This seems reasonable in that the description of a pharmaceutical product can be more precise than the description of a process and, as with all patents, it is easier to obtain evidence of infringement of product patents. One also sees the expected higher ρ indicating a longer development and commercialization period. The differences between periods are considered further below.

In short, the model estimates show that patents covering computer innovations are characterized by a moderately long development period with relatively low probability of attaining high values. However, those patents which do turn out to be valuable tend to retain their value for a longer period. In addition, while still high, the probability of successfully defending such patents is relatively low. Textile patents are similar to the computer patents with even lower potential and longer gestation periods. They are more easily defended and like computer patents tend to retain their value over time. Engine patents have high probabilities of being valuable and patentees learn quickly about their prospects. However, of the technology areas considered here, they are the most difficult to defend against challenges and they lose value rapidly over time. Pharmaceutical patents in the first part of the period had very high potential, quite low levels of depreciation, and were exploited quickly. However, these patents were relatively difficult to defend. In the second part of the period, the potential of pharmaceutical patents as whole fell relative to earlier pharmaceutical patents although it remained high relative to patents in other areas. The time lag before patents begin to generate returns grew dramatically. However, patentees could be almost certain to be able to defend their rights against challengers.

Before turning to the value of protection implied by these parameter estimates, Figures 1.a to 1.d give some idea about the fit of the model. The figures show an average of the sample and simulated hazard proportions weighted by the sample $n^g(a)$. (The simulation size is three times the sample size). The simulated hazards track the sample hazards quite closely in the first three groups. It is particularly striking that for the one group which displays a distinct fall in observed hazards in the initial ages (textiles) the model picks up the fall while, at the same time, the model hazards do not have this form for the other groups. The model also captures the rather sharp increase in the hazard in the final age.

The evolution of returns over age and the obsolescence rates implied by the parameter estimates generate two first-order stochastic dominance relationships which can be tested for consistency with non-parametric evidence. The estimated distribution of returns of first period pharmaceuticals stochastically dominates that of engines in every age and pharmaceuticals have a lower obsolescence rate. Thus the renewal curve for first period pharmaceuticals should lie everywhere on or above that of the engine group. Similarly, the estimates imply that second period pharmaceuticals dominate textiles. Both of these renewal curve relationships implied by the parameter estimates are seen in the data.

Constancy of the parameters over time in the three non-pharmaceutical groups is perhaps the most unhappy assumption of the estimations presented here since the data encompasses patents from a thirty-year period. Whether the implied stability assumption is appropriate is difficult to ascertain without actually re-estimating the model for shorter periods or with year or trend differences in some of the parameters. Since cohort differences in the cost schedule are one of the two identifying sources of variation (age differences within each cost schedule being the other) in the data, there is a constraint to how many distinctions may be drawn in this dimension. One might expect that a rise in the obsolescence rate or a shift downwards in the learning distribution over time, for instance, would be accompanied by a rise in the vector of hazard rates. A look at the residuals, $\underline{h}_N - \underline{\eta}_N(\hat{\underline{\omega}}_N)$, did not reveal any noticeable patterns over time. Since the model is very non-linear, with $\underline{\eta}_N(\hat{\underline{\omega}}_N)$ the result of several potentially offsetting processes, this observation only suggests that there is no obvious misspecification from the absence of a time dimension.

Table 4 presents characteristics of the value distributions implied by the parameter estimates. All of the distributions are very skewed - at most 16% of total value (calculated as the discounted lifetime stream of returns less annual costs) accrues to the bottom 50% of patents. However, the technology groups differ substantially in mean value per patent[19]. As one might expect, textile patents are least valuable, with a mean of 17,486 DM and a median value of just 8,514. Also, as one would expect from survey data (see, e.g.,

---

[19] The only previous model estimates which are comparable to these are in Pakes (1986) where he finds a mean value per German patent of 16,169 1980 US dollars (46,608 1975 DM) during the period 1952-72 using aggregated data. If patent values are increasing, as indicated by studies using a deterministic model (Schankerman, 1991, Pakes and Schankerman, 1986) then the estimates here should be higher. On the other hand, by using hazard rates the estimates here are not biased by the sample selection induced by the granting procedure so would tend to be lower. It is encouraging that the estimates are of the same order of magnitude.

Table 4

Value Distributions for Cohort 1975[1]

Percentile	Computers Value	Computers Cum % of Total	Textiles Value	Textiles Cum % of Total
50%	13,050 DM	9.7 %	8,514 DM	7.6 %
	(1,295)		(2,002)	
75	31,967	32.2	23,268	28.9
	(2,881)		(4,907)	
90	59,706	59.8	46,258	56.9
	(5,225)		(9,549)	
95	83,327	74.6	66,064	72.3
	(7,118)		(13,235)	
99	143,794	92.3	119,345	91.6
	(11,793)		(22,968)	
99.9	245,614	98.9	191,038	98.8
	(19,944)		(34,436)	
Mean Value	23,495		17,486	
	(2,147)		(3,689)	

Engines

Percentile	Value	Cum % of Total
50%	33,169 DM	15.4 %
	(3,314)	
75	66,402	39.4
	(6,126)	
90	114,954	65.3
	(10,215)	
95	153,731	78.5
	(13,031)	
99	248,977	93.7
	(20,522)	
99.9	433,001	99.0
	(34,979)	
Mean Value	49,728	
	(4,573)	

Table 4 continued

Pharmaceuticals

Percentile	Period 1 (Cohorts 1953-66) Value	Cum % of Total	Period 2 (Cohorts 1967-80) Value	Cum % of Total
50%	50,434 (10,534)	14.3 %	10,552 DM (1,355)	4.6%
75	106,364 (19,223)	37.0	35,048 (3,261)	23.6
90	193,651 (31,221)	63.7	77,660 (6,622)	51.9
95	268,052 (42,265)	76.9	115,493 (9,434)	68.9
99	449,758 (63,531)	93.2	210,096 (16,742)	90.3
99.9	682,514 (82,252)	99.0	332,383 (26,158)	98.6
Mean Value	81,643 (14,339)		27,450 (2,550)	
Percentage of patents receiving zero or negative value:	.01%		21.7%	

Notes:
1) Values are net of annual renewal and administration fees, as well as 476 DM for the application, examination and publication costs faced by the 1975 cohort (assuming examination at age 7 and publication at age 9). Calculations use 15000 simulation draws. All values are in 1975 Deutschmarks.

2) Estimated standard errors for the value percentile (vper) estimates due to error in $\hat{\underline{\omega}}_N$ are calculated using a Taylor approximation: $vper(\hat{\underline{\omega}}_N) \approx vper(\underline{\omega}_0) + \Gamma(\underline{\omega}_0)'(\hat{\underline{\omega}}_N - \underline{\omega}_0)$. The unknown gradient matrices $\Gamma(\underline{\omega}_0)$ are approximated with central finite difference gradients calculated at the point $\hat{\underline{\omega}}_N$. Thus $\hat{Var}(vper(\hat{\underline{\omega}}_N)) = \hat{\Gamma}(\hat{\underline{\omega}}_N)'(\hat{Var}(\hat{\underline{\omega}}_N))\hat{\Gamma}(\hat{\underline{\omega}}_N)$. Similarly for the estimated standard errors of the mean values.

Levin, *et al.* , 1987, and Mansfield, 1986), pharmaceutical patents generate a relatively high level of returns[20]. Considering the two periods, although the mean value is lower in the second period, it is very skewed and those patents which do turn out to be valuable generate quite high returns. It is particularly interesting to note the difference between the two period estimates in the number of patents which generate zero returns (or negative net of costs). Less then 1% were valueless before 1966 with a jump to 22% in the second period. This is suggestive of possible differences between patenting strategies towards process and product innovations. First, it may be that with product innovations, applications are made earlier, perhaps because testing prevents secrecy. They also may be patented more comprehensively, since secrecy is not an option and patent infringements are easier to detect. Both earlier and more comprehensive patenting of product innovations would lead to the protection of a larger proportion of innovations which end up not

---

[20] The returns generated by the creation of monopoly rights in a good depend crucially on the commercial environment. In the pharmaceutical industry in particular, price regulation may be an important impediment to fully exploiting the monopoly position and hence lower patent values where such regulation is binding. For example, Schankerman (1991) found that pharmaceutical patents in France were relatively low valued relative to those from other technology groups. France, during the period, had strictly regulated prices which were lower than other Western European countries. In Germany, on the contrary, prices during the 1953-88 period were largely unregulated and higher than in most other Western European countries (report by the Economists Advisory Group to the EEC Commission, see Scrips, 1988).

being worth much to protect. It is also interesting to note that while the mean value per patent after 1967 is only a third of that in the earlier years, because of the big increase in numbers, the estimated total annual value of patent protection going to the pharmaceutical industry grew by over 50% between 1960 and 1975. These value estimates measure the value of patent protection after secrecy and perhaps other alternative forms of protection are diminished and thus conclusions drawn about how valuable the patent system is to different industries need to take this into account. For example, while pharmaceutical product patents may have lower value than process patents, the fact that secrecy is a viable alternative for protecting the latter may mean that product innovations benefit more from the patent system.

The value distributions presented in Table 4 indicate that the value of patents varies greatly not only across but also within technology groups. From the point of view of using patents as a proxy for innovation either as a successful output of the R&D process or as an input into productivity and growth, unless this variation is accommodated, patents counts make a very noisy indicator. Using the parameter estimates from the model, one can determine the average value for patents which (i) drop in a given age or (ii) renew in a given age. If one knows the age at which each patent in a group is dropped, then one can weight them by the average value of those dropping in that age and remove much of the noise in the use of counts in measuring patent value. The extent of improvement depends on how much of the variation in value is between patents dropping in different ages and how much is among patents within each age. Table 5 shows that 39 to 56 percent of the variation in patent value is between ages. Further, 7 (engines) to 22% (pharmaceuticals) of the within age variation is among those which are never dropped. For a group of patents which all drop, even more of the variance is between age. Thus weighting groups of patents of a given technology according to drop out ages can remove a substantial part of the variance in their estimated values.

Part I of Table 5 presents the mean value of patents dropped in the indicated age as a ratio of the mean value of those dropped in age 3. Often, interest will be on groups of innovations, and patents, which are less than 20 years old. Clearly in such a case it is not possible to weight by drop out ages since most of them will have yet to drop out. Part II of Table 5 presents the ratio of the mean value of those renewed at the given age to the mean value of those dropped in age 3. This information could be used to construct weights for valuing an existing stock of patents still in force.

The weights found in Part I of Table 5 are similar to weights constructed from estimates of other patent renewal models. Schankerman (1991) presents mean value weights for four technology areas constructed from estimates of a deterministic model using French renewal data. He finds a similar linear or quadratic increasing trend in the weights with a discontinuous jump at the end.

Table 5[1]

Value-Based Weighting Schemes for Patent Stocks[1]

Analysis of Variance in Patent Value

	Computers	Textiles	Engines	Pharmaceuticals (Period 2)
% within age[2]	44%	47%	61%	52%
% between age	56	53	39	48

I. After Death of Cohort

Mean value of patents dropped in the indicated age relative to those dropped in age 3.

Age	Computers	Textiles	Engines	Pharmaceuticals (Period2)
3	1.00	1.00	1.00	1.00
4	2.90	0.51	1.46	4.73
5	4.33	1.36	1.82	6.55
6	5.26	2.20	2.35	8.51
7	6.16	2.47	2.49	8.70
8	6.44	2.96	3.02	10.62
9	7.50	3.47	3.12	11.87
10	7.71	3.16	3.11	10.92
11	8.15	3.63	3.37	11.65
12	8.89	4.12	3.64	13.56
13	10.78	4.51	3.56	13.06
14	11.65	5.10	3.83	14.69
15	13.31	5.61	4.07	14.76
16	14.39	6.32	4.31	17.07
17	16.98	8.16	5.24	21.15
18	19.71	8.80	5.64	24.07
19	24.37	10.95	6.41	33.06
20	26.25	10.38	6.44	30.38
21	39.81	16.66	10.53	43.98

Table 5 continued

II. Before Death of the Cohort

Mean value of patents renewed in the indicated age relative to those dropped in age 3.

Age	Computers	Textiles	Engines	Pharmaceuticals (Period2)
3	4.26	1.44	1.51	6.36
4	5.67	2.23	1.97	8.36
5	6.93	2.87	2.39	10.18
6	8.12	3.41	2.75	11.82
7	9.28	3.95	3.11	13.47
8	10.44	4.48	3.40	14.95
9	11.61	4.98	3.70	16.40
10	12.95	5.61	4.02	18.04
11	14.39	6.28	4.34	19.87
12	16.13	7.06	4.65	21.66
13	17.88	7.89	5.06	24.03
14	20.00	8.78	5.54	26.45
15	22.26	9.87	6.10	29.40
16	24.94	11.19	6.78	33.02
15	28.06	12.25	7.45	36.19
18	31.46	13.70	8.18	40.34
19	34.00	14.56	8.84	41.80
20	39.81	16.66	10.53	43.98

Mean Value of Those Dropped in Age 3 (1975 DM)

	Computers	Textiles	Engines	Pharmaceuticals (Period2)
	2688.99	5269.45	18281.83	2265.83

Notes:

1) Calculations from a simulation of size 15000.
2) The total variance is partitioned into within and between components using the identity:

$$\Sigma_c\Sigma_a(\eta^c_r(a) - \bar{\eta}^{\cdot}_r(\cdot))^2 = \Sigma_c\Sigma_a(\bar{\eta}^c_r(\cdot) - \bar{\eta}^{\cdot}_r(\cdot))^2 + \Sigma_c\Sigma_a(\bar{\eta}^c_r(a) - \bar{\eta}^c_r(\cdot))^2$$

where . indicates the dimension(s) over which a mean is taken, $a=5,\ldots,a_c$ and $c=53,\ldots,80$.

The high mean value for the final group is more pronounced because his model does not allow for obsolescence. In the model here, those patents which are dropped in each age due to obsolescence come from throughout the distribution, increasing the mean value in each age. Pakes and Simpson (1989) present comparable weights for the Pakes (1986) stochastic model in Pakes (1986) which was estimated using aggregate German data. Again there is a fairly smooth trend in the weights until the final category where the mean value is substantially higher. Pakes' model allows for obsolescence but its estimated value is effectively zero so again the jump in the final age is more pronounced than those seen in Table 5.[21]

It is interesting to compare the mean value of patents granted in 1975 found by weighting with the simulated mean values for 1975 found in Table 2. These are presented in the first rows of Parts I and II of Table 6. The estimated mean value of granted patents is much higher than the simulated mean value. The

---

[21] When using this type of weighting scheme it is important to remember the maintained hypothesis on which they are based - that all patents under consideration are draws from the same population. If one finds that weighted patents from two groups identifiable by some characteristic yield substantially different mean values, it suggests that they may, in fact, be draws from distinct populations making the weights inappropriate. For example, it would not be appropriate to use the weights presented here to look for trends in patent value over time since the model assumes constancy over time.

difference comes from the sample selection effect of granting. The simulation calculations start with a group of patents from age 1 while granted patents is more select in not including patents which drop before granting. The simulated value may be a more accurate reflection of the value of patenting because returns are received by patent applicants even if the application is never granted. Thus the simulated mean values may be used as a way to assess the value of those which never appear in the data as granted patents. Since considerably more patents are applied for than ever get granted, this allows one to derive information about a much larger number of innovations.[22]

One way to think of patent protection as an incentive mechanism is to consider it as an implicit subsidy to R&D. Like a direct subsidy of R&D, patent protection increases the expected return from an investment in R&D. Unlike a subsidy, however, patent protection increases the variance in returns since the implicit subsidy is only 'paid' on those innovations which prove to be successful. Assuming risk aversion, a given increase in expected returns due to patent protection is equivalent to a somewhat lower direct subsidy to R&D.

---

[22] Total simulated value is somewhat over stated since some applications fail to be granted so the 'protection' that they offer is not effective. (Compensation for infringements requires successful examination.)

Table 6

Estimates of Value and Relationship to R&D Expenditure
(1975 DM)

I. Simulated Population - Cohort 1975

	Computers	Textiles	Engines	Pharmaceuticals
Mean Value[1]	23,495	17,486	49,728	27,450
Applications	1,172	1,251	1,069	1,594
Total Value (thousands)	27,536	21,875	53,159	43,755
R&D (millions)	223	29	461	288
R&D/Applications (thousands)	190	23	431	181
Total Value/ R&D	12.3%	75.4%	11.5%	15.2%

II. Granted Population - Cohort 1975

	Computers	Textiles	Engines	Pharmaceuticals
Mean Value	41,945	29,699	74,909	53,764
Granted Patents	553	374	351	364
Total Value (thousands)	23,196	11,107	26,293	19,570
R&D/Grants (thousands)	403	78	1,313	791
Total Value/ R&D	10.4%	38.3%	5.7%	6.8%

Notes:
1) Simulated mean values are taken from Table 2. Number of applications is inferred from the number of grants in nationality/technology groups and their respective granting proportions. In Part II, the mean value of granted patents uses the weighting scheme of Table 6 applied to the sample observations.

Some average calculations of the relationship between the value created by a patent and the R&D costs of the underlying innovations are provided in Table 6. Focus is on Part I of the table since, as note above, the simulated values are more informative. The estimated implicit subsidy rates created by the patent system in each technology group are found in the last line of Parts I and II[23]. They are similar, around 10-15%, for all but textiles which is clearly different, receiving an implicit subsidy of 75%. These figures seem plausible and the lower ones are similar to those calculated for France (Schankerman, 1991).

In addition to the approximative nature of the R&D figures (details in Appendix VI), there are several points which should be borne in mind in interpreting the entries in Table 6. First, t, developing a demand for a new innovative product (i.e., product vs. brand advertising) create a public good of the same sort created by R&D expenditures and so should be included in the costs of

---

[23] These figures assume no lag between R&D expenditure and patent application. Surprisingly, very short lags have been found empirically (ie, Hall, Griliches and Hausman, 1986, for econometric results, Pakes and Schankerman, 1984, for survey evidence). To make the implicit subsidy comparable to a direct subsidy paid at the time of expenditure, the returns to patent protection should be discounted by the number of years between R&D expense and patent application.

innovation[24]. Also, the R&D figure does not include R&D done by private individuals. A survey conducted in 1975 (Oppenlander, 1977) found that 94% of all sampled patents were to protect innovations by employees and one would expect private innovation to be less R&D intensive. On the other hand, the importance of private R&D might vary by industry and one could argue that the reason for a high apparent subsidy rate in textiles is simply that relevant R&D is not captured in the statistics. If one were to assume the same response to subsidy in textiles as in the other technologies, the implied R&D is on the order of 140 million DM rather than the 29 million captured in the data.

A second point, and perhaps with most relevance to textiles, is that some innovations are not the result of R&D investments. Returns generated by patents on such innovations are not an implicit subsidy to R&D. The appropriate figure for calculating the implicit subsidy to investment in R&D would be the total value *of patents covering innovations resulting from such investment.* Not adjusting for this factor again would lead to over-estimates of the implicit subsidy.

A similar and more important point is that not all

---

[24] A study by the Expert Committee of the German Federal Ministry of Research and Technology (quoted in Beier, 1977) found that the government had difficulty finding licensees for freely available government-financed research results because without exclusive licences backed by patent protection, firms were unwilling to invest in final development and marketing.

innovations resulting from R&D are patented and the propensity to patent differs across technologies. For instance in a large survey of German firms in many technology areas (Oppenlander, 1977), respondents estimated that only 10-20% of their innovations were patented. On the other hand, pharmaceutical data indicates a propensity to patent new products of close to 100%. Since the R&D figure in the table is a total, it is larger than the R&D going to produce the innovations receiving patent protection. This means that, to the extent that patenting is less than comprehensive, the R&D per patent is over-estimated. It is likely that this factor explains a large part of the difference between the R&D per engine patent and that say, for a pharmaceutical patent. This consideration does not, however, substantially effect the interpretation of the implicit subsidy rate. Patent protection does not confer a subsidy to R&D which is *known* to be directed towards innovations which will not be patented and it confers a higher than indicated subsidy on that which is *known* to be directed toward patented, or patentable, innovations.

Concluding Remarks

To date the estimation of renewal models is the only source of large sample quantified information regarding the value of the protection offered by patent systems and how it varies across different technology areas. This information is important for understanding and formulating policy regarding intellectual property.

In addition to providing such information about an existing patent system, model estimates can also be used to simulate various policy changes in order to see their impact on the value of patent protection. For example, one could calculate the private benefit of increases in maximum patent life. One could determine the impact of changes in the fee schedule on both the value of patent protection and on renewal behavior. Simulating the impact of changes in policy regarding 'scope' (also discussed under the headings 'breadth,' 'height,' and 'novelty' - i.e. the area of product/technology space covered by a single patent) is more problematic. While it is straightforward to associate this concept with model parameters, the concept does not have an obvious metric to measure change nor is it clear what the corresponding change in parameters would be. Furthermore, contrary to what is sometimes claimed[25], the impact of a change in breadth on the value of patents does not correspond to the impact on incentives because a change in breadth will also alter the *number* of patents obtained for any level of R&D. In other words, two different policy combinations of breadth and length which yield the same average patent value do not yield the same reward to R&D expenditure.

The value distributions of patent protection estimated for the different technologies may be used in econometric studies requiring a quantitative measure of innovation. Patent counts are a very noisy

---

[25] For example, Gilbert and Shapiro (1989) and Klemperer (1989).

indicator. The advantage of exploiting the renewal data is that it allows one to give relative value weightings to individual patents based on an observable characteristic (drop out age). A drawback is that in principle this would first require the estimation of an appropriate renewal model. However, the weighting scheme derived here, and the related work using renewal models discussed above, suggests that a two parameter weighting function - an age trend and free final weight - allows a good approximation to the true series of weights. These parameters could be estimated as ancillary output of the estimation of the model of interest, improving on count data without requiring the estimation of a full patent renewal model.

References

Ashenfelter, O. (1990) "Lawyers as Agents of the Devil in a Prisoners' Dilemma Game," Working Paper no. 270. Industrial Relations Section, Princeton University.

Austin, D. (1993) "Estimating the Effects of Rival Patents in Biotechnology: An Event Study," mimeo, Dept. of Economics, University of California, Berkeley.

Beier, F-K (1980) "The Patent System and Technological, Economic and Social Progress," International Review of International Property and Copyright Law. Vol. 11, no. 5, pp. 564-84.

Berkenfeld, E. (1967) "Germany, Federal Republic of" in Hobart N. Durham (ed.) International Patent Litigation. The Bureau of International Affairs, Inc. Washington, D.C.

Bohlig, Manfred (1987) "Germany", in Andre Boujou (ed.) Patent Infringement Litigation Costs - a Practical Worldwide Survey. Chapter 7. Longman: London.

Cockburn, I. and Z. Griliches (1988) "Industry Effects and Appropriability Measures in the Stock Market's Valuation of R&D and Patents," American Economic Association Papers and Proceedings. Vol. 78, no. 2, pp. 419-423.

Gilbert, R. and C. Shapiro (1989) "Optimal Patent Length and Breadth," The Rand Journal of Economics. Vol 21, no. 1.

Gill, P., Murray, W. and M. Wright (1981) Practical Optimization. Academic Press Limited: London.

Griliches, Zvi (1990) "Patent Statistics as Economic Indicators: A Survey", Journal of Economic Literature. Vol. XXVIII (December, 1990). pp. 1661-1707.

Hall, B. H., Griliches, Z. and Hausman, J. (1986) "Patents and R&D: Is There a Lag?" International Economic Review. Vol. 27, no. 2, pp. 265-83.

Klemperer, Paul (1989) "How Broad Should the Scope of Patent Protection Be?" The Rand Journal of Economics. Vol. 21, no. 1.

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

Lanjouw, J. O. (1992a) The Private Value of Patent Rights: A Dynamic Programming and Game Theoretic Analysis of West German Patent Renewal Data, 1953-1980. Ph.D. Thesis, London School of Economics and Political Science.

Lanjouw, J. O. (1992b) "Under Threat: Potential Competition, Litigation and the Private Value of Patent Protection," Economics of Industry Group Working Paper no. EI/6. STICERD. London School of Economics.

Levin, R., Klevorick, A. Nelson, R. and S. Winter (1987) "Appropriating the Returns from Industrial Research and Development", Brookings Papers on Economic Activity, 3, pp. 783-820.

Mansfield, E. (1986) "Patents and Innovation: An Empirical Study", Management Science, vol. 32, no. 2, pp. 173-181.

Mansfield, E., Schwartz, M. and Wagner, S. (1981) "Imitation Costs and Patents: An Empirical Study," The Economic Journal. Vol 91, no. 364, pp. 907-918.

McFadden, D. (1989) "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration" Econometrica. Vol 57, no. 5, pp. 995-1026.

Oppenlander, Karl-Heinrich (1977) "Patent Policies and Technological Progress in Germany," International Review of Intellectual Property and Copyright Law. Vol. 8, no. 2, pp.97-112.

Pakes, A. (1985) "On Patents, R&D, and the Stock Market Rate of Return," Journal of Political Economy.Vol 93, no. 2, pp. 390-409.

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

Pakes, A. and D. Pollard (1989) "Simulation and the Asymptotics of Optimization Estimators" Econometrica. Vol. 57, no. 5. pp. 1027-58.

Pakes, A. and M. Shankerman (1984) "The Rate of Obsolescence of Knowledge, Research Gestation Lags, and the Private Rate of Return to Research Resources" in Zvi Griliches (ed.) Patents, R&D and Productivity. Chicago: The University of Chicago Press.

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

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

Scrip: World Pharmaceutical News (1988) May 13, no. 1308.

Stauder, Dieter (1983) "The Practical Significance of Infringement and Revocation Proceedings in the Federal Republic of Germany, France, Italy and the UK. Results of a Statistical and Empirical Study". International Review of Intellectual Property and Copyright Law. Vol. 14, no. 6, pp. 793-810.

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.

Appendix I
Characteristics of the Data

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

Mean Sample Size per Cohort[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 cell samples are equivalent to the entire population of granted patents for every cohort covered.

</ref_section>

Appendix II
Patenting Procedure, Fees and Legal Rights

Stage	Fees[26]	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 court 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 court challenges		
Granting	Annual renewal fees	

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

Appendix III

Derivation of Estimating Equations

Parts I and II, which characterize expected value and derive a set of simplifying boundary equations, are based on similar derivations in Pakes (1984). Part III incorporates the litigation effects to obtain the equations used in the optimization routine.

Part I. Characterization of Expected Value

Returns in age a+1 are equal to max{ δr_a, z} where z is an annual draw from an exponential distribution with cdf:

$$Q_{a+1}(z) = 1\text{-}\exp\{-(z/\sigma_{a+1}+\gamma)\}$$

Because of the discontinuity in the cdf of r_{a+1}, the characterization of EV(a+1) differs according to whether $\delta r_a <> \bar{r}_{a+1}$. If $\delta r_a > \bar{r}_{a+1}$ then, unless obsolescence occurs, the value in a+1 will be strictly positive with r_{a+1} equal either δr_a or z. If $\delta r_a \leq \bar{r}_{a+1}$ then whether the value in a+1 is greater than zero depends on the realization of z, the draw in age a+1. For example, consider the final age, which throughout the discussion is treated as 20.

$$E[V(20)\,|\,r_{19},.] = \qquad \text{(III.1)}$$

a) If $r_{19} \leq \bar{r}_{20}/\delta$

$$\int_{\bar{r}_{20}}^{\infty} (z-c_{20})\, dQ_{20}(z)$$

$$= h^0_{19}$$

b) if $r_{19} > \bar{r}_{20}/\delta$

$$\int_{-\gamma\sigma_{20}}^{\delta r_{19}} (\delta r_{19}-c_{20})\, dQ_{20}(z)$$

$$+ \int_{\delta r_{19}}^{\infty} (z-c_{20})\, dQ_{20}(z)$$

$$= h^0_{19} + h^1_{19}(r_{19})$$

As indicated, $h^0_{19}(r_{19}) = E[V(20)\,|\,r_{19} \leq \bar{r}_{20}/\delta]$ is the expected value in age 20 when r_{19} falls in the first range and $h^1_{19}(r_{19})$ is the incremental increase in expected value when r_{19} falls in the second range $r_{19} > \bar{r}_{20}/\delta$.

Working backwards, the same distinction occurs for age 19:

$$E[V(19)|r_{18},.] = \tag{III.2}$$

a) if $r_{18} \leq \bar{r}_{19}/\delta$

$$\int_{\bar{r}_{19}}^{\bar{r}_{20}/\delta} [z-c_{19}+\beta\theta h_{19}^{0}]\ dQ_{19}(z)$$

$$\int_{\bar{r}_{20}/\delta}^{\infty} [z-c_{19}+\beta\theta(h^{0}+h^{1})_{19}(z)]\ dQ_{19}(z)$$

$$= h_{18}^{0}$$

b) if $\bar{r}_{19}/\delta \leq r_{18} \leq \bar{r}_{20}/\delta^2$

$$\int_{-\sigma_{19}\gamma}^{\delta r_{18}} [\delta r_{18}-c_{19}+\beta\theta h_{19}^{0}(\delta r_{18})]\ dQ_{19}(z)$$

$$+ \int_{\delta r_{18}}^{\bar{r}_{20}/\delta} [z-c_{19}+\beta\theta h_{19}^{0}(z)]\ dQ_{19}(z)$$

$$+ \int_{\bar{r}_{20}/\delta}^{\infty} [z-c_{19}+\beta\theta(h^{0}+h^{1})_{19}(z)]\ dQ_{19}(z)$$

$$= (h^{0}+h^{1})_{18}(r_{18})$$

c) if $r_{18} > \bar{r}_{20}/\delta^2$

$$\int_{-\gamma r_{19}}^{\delta r_{18}} [\delta r_{18}-c_{19}+\beta\theta(h^{0}+h^{1})_{19}(\delta r_{18})]\ dQ_{19}(z)$$

$$+ \int_{\delta r_{18}}^{\infty} [z-c_{19}+\beta\theta(h^{0}+h^{1})_{19}(z)]\ dQ_{19}(z)$$

$$= (h^{0}+h^{1}+h^{2})_{18}(r_{18})$$

where the ranges defined are non-empty by virtue of the fact that $\bar{r}_a$ is increasing in a.

In general define $h_a^0(a)$ as the expected value in a+1 when the returns in a fall in the first interval, $r_a \leq \bar{r}_{a+1}/\delta$, and $h_a^v(r_a)$, v = 1,...,(20-a), as the incremental increase in the expected value in a+1 when $r_a \geq \bar{r}_{a+v}/\delta^v$. Because no returns are received the terminal age, EV(21)=0 and $h^v_{20} = 0$ for all v.

h_a^0 is composed of two parts:

i) the expected net returns in a+1 (III.3)

$$\int_{\bar{r}_{a+1}}^{\infty} (z-c_{a+1})\ dQ_{a+1}(z)$$

ii) the expected value in age a+2. This is the sum of the incremental increases in the expected value at a+2 integrated over the relevant ranges of r_{a+1}. (III.4)

$$\beta\theta\sum_{v=0}^{20-(a+1)} \int_{\bar{r}_{a+v+1}/\delta^{v}}^{\infty} h_{a+1}^{v}(z)\ dQ_{a+1}(z).$$

or, to simplify notation, $\beta\theta\Sigma_{v=0}^{20-(a+1)} H_{a+1}^{v}$. Note that h_a^0 is independent of r_a.
For v=1, (III.5)

$$h_a^1(r_a) = [Q_{a+1}(\bar{r}_{a+1})][\delta r_a - c_{a+1} + \beta\theta(h_{a+1}^0)]$$

$$+\beta\theta\int_{\bar{r}_{a+1}}^{\delta r_a}[\delta r_a - z]dQ_{a+1}(z).$$

For $1 < v \leq 20$-a, (III.6)

$$h_a^v(r_a) = [Q_{a+1}(\bar{r}_{a+v}/\delta^{v-1})][\beta\theta h_{a+1}^{v-1}(\delta r_a)]$$

$$+\beta\theta\int_{\bar{r}_{a+v}/\delta^{v-1}}^{\delta r_a}[h_{a+1}^{v-1}(\delta r_a) - h_{a+1}^{v-1}(z)]dQ_{a+1}(z)$$

The fact that $\bar{r}_a$ is increasing in a and that $\delta \leq 1$, together insure that $\bar{r}_a \leq \bar{r}_{a+1}/\delta$. Thus $\bar{r}_a$ is the solution to

$$r_a - c_a + \beta\theta h_a^0 = 0. \qquad \text{(III.7)}$$

Because h_{a+1}^0 is a function of known c, the parameters ω and only future $\bar{r}$'s, it is theoretically possible to analytically solve the required integrals in a recursive fashion. However, the equations for successively earlier $\bar{r}$'s quickly become intractable even with the aid of mathematical software. Pakes (1984, Appendix 3) notes that the integrals H_a^v may be approximated by the following, more simply calculated, boundary functions.

Part II - Boundary Equations

First note that for every a, $h^v_a(x)$ is positive and increasing in x for $v \geq 1$.

1) Upper Bound

From equation III.6, (III.8)

$$h_a^v(r_a) = \beta\theta Q_{a+1}(\delta r_a) h_{a+1}^{v-1}(\delta r_a) - \beta\theta \int_{\bar{r}_{a+v}/\delta^{v-1}}^{\delta r_a} h_{a+1}^{v-1}(z)\, dQ_{a+1}(z)$$

Then, (III.9)

$$h_a^v(r_a) \leq \beta\theta h_{a+1}^{v-1}(\delta r_a) - \beta\theta \int_{\bar{r}_{a+v}/\delta^{v-1}}^{\delta r_a} h_{a+1}^{v-1}(z)\, dQ_{a+1}(z)$$

$$\leq \beta\theta h_{a+1}^{v-1}(\delta r_a)$$

Where the second inequality follows from the fact that $h^{v-1}_{a+1}(z) \geq 0$ and that $\delta r_a \geq \bar{r}_{a+v}/\delta^{v-1}$ (the limits of integration) where h^v_a applies.

Substituting recursively, (III.10)

$$h_a^v(r_a) \leq (\beta\theta)^{v-1} h_{a+v-1}^1(\delta^{v-1} r_a)$$

for v=1,...,20-a.

Using this inequality in the definition of H_{a+1}^v, (III.11)

$$H_{a+1}^v \leq (\beta\theta)^{v-1} \int_{\bar{r}_{a+v+1}/\delta^v}^{\infty} h_{a+v}^1(\delta^{v-1} z)\, dQ_{a+1}(z).$$

2) Lower Bound (III.12)

Because $h_{a+1}^{v-1}(z)$ increases monotonically in z,

$$\int_{\bar{r}_{a+v}/\delta^{v-1}}^{\delta r_a} h_{a+1}^{v-1}(z)\ dQ_{a+1}(z) \le \int_{\bar{r}_{a+v}/\delta^{v-1}}^{\delta r_a} h_{a+1}^{v-1}(\delta r_a)\ dQ_{a+1}(z)$$

$$= h_{a+1}^{v-1}(\delta r_a)[Q_{a+1}(\delta r_a)-Q_{a+1}(\bar{r}_{a+v}/\delta^{v-1})]$$

Substituting this inequality into equation III.8,

(III.13)

$$h_a^v(r_a) \ge (\beta\theta)^{v-1}\kappa_a^v\ h_{a+v-1}^1(\delta^{v-1}r_a)$$

for v=2,...,20-a, where (III.14)

$$K_a^v = \prod_{i=1}^{v-1} Q_{a+i}(\bar{r}_{a+v}/\delta^{v-i})$$

This leads to the lower bound, (III.15)

$$H_{a+1}^v \ge (\beta\theta)^{v-1}\ \kappa_{a+1}^v \int_{\bar{r}_{a+v+1}/\delta^v}^{\infty} h_{a+v}^1(\delta^{v-1}z)\ dQ_{a+1}(z)$$

for v=1,...,20-a.

Using the midpoint between the upper and lower bounds, the approximate H_a^v equation is:

(III.16)

$$H_{a+1}^v \approx [(1+\kappa_{a+1}^v)/2][(\beta\theta)^{v-1}\int_{\bar{r}_{a+v+1}/\delta^v}^{\infty} h_{a+v}^1(\delta^{v-1}z)\ dQ_{a+1}(z)]$$

Part III - Estimation Equations Incorporating Litigation

The calculation of the minimum level of returns required for a decision to renew when the need to defend patent rights is introduced into the model is a reasonably straightforward alteration of the procedure described above. From Part I recall that if $\delta r_a < \bar{r}_{a+1}$ the value in a+1 is greater than 0 only for realizations of $z > \bar{r}_{a+1}$. Because owners are assumed to make optimal future decisions, any realization of $z < \bar{r}_{a+1}$ will lead to the patent being dropped and a value in a+1 of 0. Thus, the limits of integration in this case run from $\bar{r}_{a+1}$ to infinity. If $\delta r_a > \bar{r}_{a+1}$, then the value in a+1 is positive for any realization of z. In this case, the integral to find E[V(a+1)|.] is over the entire support of z. The difference here is that a decision to renew in age a when owners know that they face the possibility of litigation implies a willingness (in expectation given the information available at age a) to maintain the patent in force over the period of litigation if challenged. Thus it implies a willingness to continue under the expectation that any realizations of z will lead to renewal over the period of litigation regardless of the current level of returns. As a result, for ages a+1 and a+2, the limits of integration are over the entire support of z.

To simplify notation, define $G_a(k|r_{a-1})=\text{prob}(r_a<k|r_{a-1})$. G_a incorporates Q_a and the mass point at δr_{a-1}. The period of litigation is taken to be two years past the current age. The equation for $E[VL(a+1)|r_a,.]$ is, (III.17)

$$E[VL(a+1)|r_a,.] = \int s G_{a+1}(ds|r_a)$$

$$+\beta\int\int(x + \beta^2\theta E[V(a+3)|x,.])\ G_{a+2}(dx|s)\ G_{a+1}(ds|r_a)$$

where VL is used to denote value from the second year of an infringement suit. c_{a+1}, c_{a+2} and θ over the two-year period of litigation have been removed to enter the new decision rule explicitly below.

In terms of the h^v_a terms this means, for example, that $E[VL(20)|r_{19}]$ is calculated as in case (b) of equation III.1 regardless of the level of r_{19}. In general, when determining the expected value of renewal for any age a, the limits of integration for the following two years run from $-\gamma\sigma_{a+1}$ and $-\gamma\sigma_{a+2}$ to infinity and, as a result, the first two break points in the range of r_a ($\bar{r}_{a+1}/\delta$ and $\bar{r}_{a+2}/\delta^2$) are no longer significant.

$$E[VL(a+1)|r_a] = \tag{III.18}$$

a) If $r_a \leq \bar{r}_{a+3}/\delta^3$ $\qquad (h^0 + h^1 + h^2)_a(r_a)$

b) If $r_a > \bar{r}_{a+i}/\delta^i$ $\qquad$ as previously, $\Sigma_{v=0}{}^{i}h_a{}^v(r_a)$

for i=3,..., (20-a) with costs and θ for ages a+1 and a+2 removed.

The definitions of $h_a{}^v$ and $H_a{}^v$ remain the same so the boundary equations derived in Part II continue to be appropriate. There is one caveat. It is no longer generally true that $\bar{r}$ is non-decreasing in a, all a. Because of the preparedness effect, with some values for the model parameters it is possible to find $\bar{r}_{18} > \bar{r}_{19} > \bar{r}_{20}$. However, parameter restrictions required for consistency ensure that this situation does not arise.

Age 20

One arrives at $\bar{r}_{20}$ directly from the decision rule (2). $\bar{r}_{20}$ solves:

$$V(20, \bar{r}_{20}) = w\bar{r}_{20} - c_{20} - (1-w)[\alpha_0 + \alpha_1\bar{r}_{20}] = 0. \tag{III.19}$$

Note that $V(20, r_{20}=0)$ is strictly negative since there is no option for future ages. Hence $\bar{r}_{20}$ is strictly positive.

$$\Rightarrow \qquad \bar{r}_{20} = [c_{20} + \alpha_0(1-w)]/[w - \alpha_1(1-w)].$$

Age 19

r_{19} is defined implicitly by the following equation:

$$V(19, r_{19}) = \theta w r_{19} + \beta\theta w[h^0+h^1]_{19}(r_{19}) - c_{19} - \beta\theta c_{20} - (1-w)[\alpha_0+\alpha_1 r_{19}] = 0. \tag{III.20}$$

with

$$r_{19} = 0 \text{ iff}$$
$$V(19, r_{19} = 0) \geq 0.$$

(Only one θ appears because θ in age 21 equals one - there is no obsolescence where there are no returns.)

Age < 19

Other r's are defined implicitly by the following equation:

$$V(a, r_a) = \theta^2 w r_a + \beta\theta^2 w[h^0+h^1+h^2]_a(r_a) - c_a - \beta\theta c_{a+1} - (\beta\theta)^2 c_{a+2} - (1-w)[\alpha_0+\alpha_1 r_a] = 0. \tag{III.21}$$

with

$$r_a = 0 \text{ iff}$$
$$V(a, r_a = 0) \geq 0.$$

Because r_a enters the second term exponentially, the solutions are found numerically.

$$(h^0 + h^1 + h^2)_a(r_a) = \tag{III.22}$$
$$\delta r_a + \sigma_{a+1}[1 - Q_{a+1}(\delta r_a)]$$
$$+ \beta\{\delta^2 r_a + [1 - Q_{a+1}(\delta r_a)][\delta\sigma_{a+1} + \xi_a^v \sigma_{a+2}[1 - Q_{a+2}(\delta^2 r_a)]]$$
$$+ \sigma_{a+2} Q_{a+1}(\delta r_a)[1 - Q_{a+2}(\delta^2 r_a)]\}$$
$$+ \beta^2\theta h^0_{a+2} + \beta^2\theta \Sigma_{v=1}^{20-(a+1)} H_{a+2}^v + \beta \Sigma_{v=2}^{20-(a+1)} H_{a+1}^v.$$

where $\xi_a^v = (\sigma_{a+v})^2/(\sigma_{a+v} + \delta^v \sigma_a)$.

$h_a^0 = [r_{a+1} - c_{a+1} + \sigma_{a+1} + \beta\theta h_{a+1}^0][1 - Q_{a+1}(r_{a+1})] + \beta\theta \Sigma_{v=1}^{20-(a+1)} H_{a+1}^v$.

$Ha+1_v$ is defined as in equation III.16 where

$$\int_{\bar{r}_{a+1+v}/\delta^{v}}^{\infty} h_{a+v}^{1}(\delta^{v-1}z)\, dQ_{a+1}(z) =$$

$$[1-Q_{a+1}(\bar{r}_{a+1+v}/\delta^{v})]\{\, \bar{r}_{a+1+v}+ \delta^{v}\sigma_{a+1}$$

$$-[1-Q_{a+1+v}(\bar{r}_{a+1+v})][\sigma_{a+v} +\bar{r}_{a+1+v}- \xi_{a+1}^{v}]\, \}$$

$$+(\beta\theta h_{a+1+v}^{0}-c_{a+1+v})\, [Q_{a+1+v}(\bar{r}_{a+1+v})][1-Q_{a+1}(\bar{r}_{a+1+v}/\delta^{v})]$$

and K_{a+1}^{v} is as defined above (eqn III.14). Note that the first $\bar{r}$ to enter the decision rule for age a is $\bar{r}_{a+3}$ which is a result of the fact that returns less than $\bar{r}_{a+1}$ and $\bar{r}_{a+2}$ no longer lead to non-renewal.

Appendix IV

Distribution of $\sqrt{N}\, G_N(\omega_0)$.

Decompose the objective function:

$$\sqrt{N}\underline{G}_N(.) = \sqrt{N}(\underline{h}_N - \underline{\eta}(.)) - \sqrt{N}(\underline{\eta}_N(.) - \underline{\eta}(.)) \qquad \text{(IV.1)}$$

Let $(N/n^c(a))$ denote the m-vector with elements equal to the sample size divided by the number of patents still in force, $n^c(a)$, at the beginning of the appropriate cohort/age. Take the portion of the m-vectors relating to a particular cohort, c, and consider the variance of the observed hazard for age a, $h_N(a)$. The 0/1 decision to renew a patent is independent of decisions made about other patents and therefore has a binomial sampling distribution. For a given n(a), the variance is

$$\text{var}(\sqrt{N}\, h_N(a)) = \eta(a)(1 - \eta(a))\, [N/n(a)]. \qquad \text{(IV.2)}$$

N/n(a) is a random variable. However,

$$N/n(a) \rightarrow 1 \,/\, \{p_c\, \Sigma_{k=1}^{a-1}\, [P(grtk)\, \Pi_{j=k+1}^{a-1}\, (1 - \eta(k;\, \underline{\omega}_0))]\, \} = 1/\lambda^1(a) \qquad \text{(IV.3)}$$

where p_c is the constant proportion of the total sample which is from cohort c and P(grtk) is the probability of being granted in age k. Therefore, the central limit theorem ensures that the first term on the RHS of equation IV.1 converges in distribution to an m-variate normal

$$\sqrt{N}(\underline{h}_N - \underline{\eta}(\underline{\omega}_0)) \rightarrow N_m(\, \underline{0},\, \text{diag}[\, \underline{\eta}(\underline{\omega}_0)(1-\underline{\eta}(\underline{\omega}_0))/\, p_c\, \text{diag}[\underline{\lambda}^1] \qquad \text{(IV.4)}$$

$$= N_m(\, \underline{0},\, V_1)$$

where $\underline{\lambda}$ is an m-vector with elements as defined in equation (IV.3). The distribution of the second term differs in that in the simulation there is no randomness with respect to the proportion which drop due to obsolescence (see Appendix V). $\underline{\eta}$ can be decomposed into non-renewals due to returns being less than $\bar{r}$, which is denoted $\underline{\bar{\eta}}$, and those due to obsolescence.

$$\underline{\eta}_N(.) = \underline{\bar{\eta}}_N(.) + (1 - \underline{\bar{\eta}}_N(.))(1-\theta). \qquad \text{(IV.5)}$$

At $\underline{\omega}_0$, the simulation draws are equivalent to random sampling *with respect to the distribution of returns*, so the same derivation of the distribution applies. Taking any cohort, c, and age a,

$$(N/n_{sim}(a)) \rightarrow 1/rp_c\, \Pi_{j=1}^{a-1}(1 - \bar{\eta}(j;\, \underline{\omega}_0)) \qquad \text{(IV.6)}$$

So,

$$\sqrt{N}(\underline{\bar{\eta}}_N(\underline{\omega}) - \underline{\bar{\eta}}(\underline{\omega}_0)) \rightarrow N_m(\, \underline{0},\, \text{diag}[\, \underline{\bar{\eta}}(\underline{\omega}_0)(1-\underline{\bar{\eta}}(\underline{\omega}_0))/\, (rp_c\, \text{diag}[\lambda^2]\,). \qquad \text{(IV.7)}$$

where $\underline{\lambda}^2$ is an m-vector with elements $(1/\theta^{a-1})\ \Pi_{j=1}^{a-1}(1 - \eta(j;\ \underline{\omega}_0))$. ($\hat{\eta}$ is removed from the denominator using equation V.5.) Thus the second term on the RHS of equation IV.1 is distributed:

$$\sqrt{N}(\underline{\eta}_N(\underline{\omega}_0) - \underline{\eta}(\underline{\omega}_0)) \rightarrow$$

$$N_m(\ \underline{0},\ (1+(1-\theta)^2)\text{diag}[\ \underline{\hat{\eta}}(\underline{\omega}_0)(1-\underline{\hat{\eta}}(\underline{\omega}_0))]\ /\ (rp_c\ \text{diag}[\underline{\lambda}^2])\ ) \tag{IV.8}$$

$$= N_m(\ \underline{0},\ V_2)$$

Since the sample and simulation draws are independent,

$$\sqrt{N}\ \underline{G}_N(\underline{\omega}_0) \rightarrow$$

$$N_m(\ \underline{0},\ V_1 + V_2) = N_m(\underline{0},\ V) \tag{IV.9}$$

which can be used in equation 4 to obtain the distribution of $\underline{\hat{\omega}}_N$. In equation 4, both Γ and V are evaluated at ω_0 which is unknown. Pakes and Pollard show that $V(\underline{\hat{\omega}}_N)$ is a consistent estimator of V. A finite difference estimator of Γ evaluated at $\underline{\hat{\omega}}_N$ is consistent as long as the finite differences are specified to shrink sufficiently in N such that $\Gamma_N \xrightarrow{p} \Gamma$.

Appendix V

Outline of the Optimization Procedure

The optimization routine was written in Gauss 386-VM programming language. A summary of the major components follows.

Objective Function

The following outlines the subprocedure to calculate the objective function.

1) Input the parameter vector, renewal fees, and hazard data.

2) Transform the matrix of random draws from a uniform distribution into one of random draws from exponential distributions parameterized by (γ,σ,ϕ) using the inverse transform. Random draws, z, from the uniform distribution are transformed by setting z equal to $Q_a(y)$ and solving for y,

$$y = -\sigma_a \ln(1-z) - \sigma\gamma.$$

For each cohort:

3) Calculate the vector of $\bar{r}$'s

The calculation of $\bar{r}_{20}$ is straightforward (Appendix III, equation III.19). The other $\bar{r}$'s must be solved using an numerical procedure. It was noted above that, with certain acceptable restrictions on the parameters, $\bar{r}$'s are non-decreasing in age, and that the expected value of renewing in any age a is non-increasing in r_a. These two results are used to find the recursive solutions for $\bar{r}$. For age 19, set $\bar{r}_{19}$ equal to $\bar{r}_{20}$ and solve the value function (eqn 2), which will be positive. Reduce $\bar{r}_{19}$ by increments[27] until either the value function becomes negative <u>or</u> $\bar{r}_{19} \leq 0$. From age 19 continue to work down, using information from the previous rounds of the recursive system.

Because the $\bar{r}$'s are positive and non-decreasing in age, as soon as an $\bar{r} = 0$ is found, end the procedure and set all $\bar{r}$'s for lower ages to zero.

4) Calculate the Simulated Hazards - $\eta_N(\omega)$

These are found first as mortality rates which are then transformed into hazard rates. The calculations are done in two stages. The first stage comprises ages 1 to t where t is the last age with $\bar{r}_a = 0$. There are no drops except for obsolescence. In age 1, returns for the simulated patents are the first column of the matrix of random draws. In age 2, depreciate these by $(1-\delta)$ and compare them element-by-element with the second column. Retain the maximum as returns for age 2, and so on.

---

[27] Specifically, $\bar{r}$ is first reduced by increments of 100. When the expected benefit function becomes negative, $\bar{r}$ is increased by increments of 10. When the function becomes positive again the procedure stops and $\bar{r}$ is set equal to its final value minus 5. The resulting estimate has a maximum error of +/- 5.

The effect of obsolescence is inferred and no simulated observations are deleted[28]. The obsolescence rate $(1-\theta)$ is the probability of becoming obsolete in a given renewal through a-1. The probability of renewing through a-1 (when the only reason to drop is obsolescence) is θ^{a-1}. Set the mortality rates in the first stage to

$$\pi^{s}_{a} = (1-\theta)*\theta^{a-1}. \tag{V.1}$$

In the second stage, where $r > 0$, find the vector of returns for age t+1 and delete the elements which are less than r_{t+1}. Count the number remaining, n_{t+1}, and again infer obsolescence without deleting observations. The number of drops in t+1 is $(N - \theta n_{t+1})$. In the second age, again find returns by taking the first n_{t+1} elements of the next column of the matrix of random draws and comparing them with those already achieved. Delete returns less than r_{t+2}. Count the number remaining, n_{t+2}. Of these, only $\theta^2 n_{t+2}$ would have remained after two periods of obsolescence so the number dropping in t+2 is $(\theta n_{t+1} - \theta^2 n_{t+2})$. Continuing, set the simulated mortality rate in age a > t to

$$\pi^{s}_{a} = \theta^{a-1}(n_{a-1} - \theta n_{a})/N. \tag{V.2}$$

If no elements remain, end the procedure and set the simulated mortality rates for the following ages to zero. Transform the simulated mortality rates into hazard rates by dividing each mortality rate by 1 minus the cumulated mortality rates for earlier ages.

Pick out the pieces of the data matrix of sample hazard rates and of numbers of patents alive at the beginning of each age, n(a), which correspond to the appropriate cohort. Calculate

$$\Sigma_{a} [n(a)/N](h_N(a) - \eta_r(a))^2 \tag{V.3}$$

Sum these over cohorts and take the square root.

<u>Computation of the Search Direction - Gradient and Hessian</u>

The direction of search is found using a quasi-Newton method. Let $\|\underline{G}_N(\underline{\omega})\|$ be denoted $f(\underline{\omega})$. For each iteration a quadratic approximation to the function is formed from the first three terms of a Taylor series expansion about the current point,

$$f(\underline{\omega}+\underline{p}) = f(\underline{\omega}) + \Gamma'(\underline{\omega})\underline{p} + \underline{p}'H(\underline{\omega})\underline{p}/2 \tag{V.4}$$

where $\underline{p}$ is a direction vector and Γ and H are gradient and Hessian matrices evaluated at the current point $\underline{\omega}$. The function at $(\underline{\omega}+\underline{p})$ is minimized when

[28]The advantage of inferring obsolescence rather than deleting simulated observations is that the effective simulation size used to construct the simulated hazards at each age is larger for a given number of calculations. Specifically, no comparisons of δr_{a-1} vs. z need to be done for elements which end up becoming obsolete. This allows much smaller simulation errors for the same calculations - particularly at higher ages.

$$\underline{p} = -\Gamma(\underline{\omega})H^{-1}(\underline{\omega}). \qquad (V.5)$$

So $\underline{p}$, estimated with approximations to Γ and H^{-1}, is used as the direction of search.

Gradient

The gradient sub-procedure calculates central finite difference gradient approximations. Perturb in turn each parameter, ω_i, in a positive and negative direction by the amount $\Delta\omega_i$ and use the gradient formula:

$$\partial f/\partial\omega_i = [f(\underline{\omega}+\Delta\omega_i) - f(\underline{\omega}-\Delta\omega_i)]/2\Delta\omega_i, \qquad (V.6)$$

where $\Delta\omega_i$ denotes either a scalar or a vector with ith element $\Delta\omega_i$ and zeros else as appropriate. This formula, which is found from a Taylor series expansion around $\underline{\omega}$, has two sources of error. First, truncation error from the fact that the second and higher order derivatives are dropped. Second, condition error from inaccuracies in computing the function values. The bound on the error from these two sources is

$$(\Delta\omega_i)2\ |f'''(\zeta)|\ /6 + \epsilon_a/\Delta\omega_i \qquad (V.7)$$

where $\zeta \in [\underline{\omega}+\Delta\omega_i, \underline{\omega}-\Delta\omega_i]$ and ϵ_a is the maximum error in the computation of $f(\underline{\omega})$. Clearly, one part of the error increases in $\Delta\omega_i$ and the other decreases, with the minimum where

$$\Delta\omega_i = d^c_i = \sqrt[3]{[3\epsilon_a/f'''(\zeta)]}. \qquad (V.8)$$

The procedure used to find optimal finite differences is suggested by Gill, Murray and Wright (1981). Two components are required, an estimate of the higher order derivative and an estimate of the error in function evaluation, ϵ_a.

Because there is no source of information about the true value of f at any point, it is only possible to estimate ϵ_a. This is done as follows:
Calculate values $f(\underline{\omega}_j)$ where $\underline{\omega}_j = \underline{\omega} + j*.00001$, j=1 to 15.
Find the differences $f(\underline{\omega}_j)-f(\underline{\omega}_{j+1})$.
Because the difference in $\underline{\omega}$ is so small, almost all of the difference in the function estimates is error. Continue taking higher order differences (differences of the previous column of differences) until they are similar in magnitude and alternate sign, which will happen given certain assumptions about the distribution of errors. Let k be the column in which this occurs and $\max\Delta f$ represent the largest difference in column k. Then ϵ_a is estimated by

$$\hat{\epsilon}_a = \max\Delta f/\sqrt{[(2k)!/(k!)^2]} \qquad (V.9)$$

Gill, Murray and Wright (1981) suggest approximating the optimal central finite difference using the following formula,

$$d^c_i = \sqrt[3]{[4\hat{\epsilon}_a/\hat{\Phi}]} \qquad (V.10)$$

where $\hat{\Phi}$ is an reasonable estimate of the second derivative calculated using the second-order difference formula:

$$\hat{\Phi}(\Delta\omega_i) = [f(\underline{\omega}+\Delta\omega_i) - 2f(\underline{\omega}) + f(\underline{\omega}-\Delta\omega_i)]/(\Delta\omega_i)^2. \qquad (V.11)$$

Once satisfactory estimates of Φ and ϵ_a are found, calculate d^c_i for each parameter. At any vector $\underline{\omega}$, the finite differences used to calculate the gradient are:

$$\Delta\omega_i = d^c_i(1 + \omega_i)/(1 + \omega_{i0}) \qquad \text{(V.12)}$$

where ω_{i0} is the value of parameter ω_i at which d^c_i was calculated. If a perturbation $+\Delta\omega_i$ causes a parameter to cross an upper bound then take the central difference step in the opposite direction:

$$\partial f/\partial\omega_i = [f(\underline{\omega}\text{-}2\Delta\omega_i) + 3f(\underline{\omega}) - 4f(\underline{\omega} - \Delta\omega_i)]/2\Delta\omega_i. \qquad \text{(V.13)}$$

and vica versa for a perturbation which crosses a lower bound. If both perturbations cause an increase in the function value then set the gradient to zero. If both perturbations cause an decrease in the function value calculate the gradient as above.

Both ϵ_a and the vector of optimal central difference intervals, $\underline{d}^c$, were calculated once for each industry. These calculations were made after the grid search and some experimentation to find a good set of parameter values at which to evaluate them. A set of example calculations is found below.

Hessian

Because it is very time consuming to evaluate the function, the hessian is not evaluated at every point. Rather, information about the curvature of the function is derived from the function and gradient evaluations as the iterations proceed. As in the estimation of Φ above, the hessian is approximated using the first term of a Taylor expansion - which requires only the change in the gradient. For the first iteration the inverse hessian is set to the identity matrix. At each subsequent iteration, calculate a new inverse hessian using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) updating procedure.

Computation of the Step Length

At each iteration the step length is found using a golden section search. This consists of bracketing the minimum along the search direction p and reducing the interval while keeping the minimum bracketed.

First calculate the minimum step length, α^{min}, based on ϵ_a:

$$\alpha^{min}\| \underline{p} \| = \epsilon_a(1 + \|\omega\|). \qquad \text{(V.14)}$$

A minimum step length is set to avoid the calculation of the function at points so near that differences will be swamped by error in the function evaluation. Take a step of .01 or α^{min}, whichever is larger, until the function begins to increase.

To reduce the interval:

Letting l be the length of the interval between ω and the last point evaluated, evaluate the function at $\underline{\omega}+.3820l$ and $\underline{\omega}+.6180l$. If at least one function value is lower, keep the highest of the two and discard the nearest endpoint. Continue. If the function value is higher at both internal points than at the endpoints, keep the endpoint with the lowest value and the nearest internal point to locate the local minimum. This avoids the best point being discarded. (Since $\underline{p}$ is a decent direction, if this occurs it indicates that a quadratic approximation is not accurate over the interval.) If the step length procedure is unsuccessful, check the function values found during the gradient estimation. If at least one of these is lower that the initial function value, go to that point and begin the next iteration. Otherwise evaluate the function at five randomly chosen points in the neighborhood. If there is no improvement, then end.

Appendix VI

Calculation of R&D Expenditure

There are two major considerations in calculating the R&D expenditure relevant to the patent values measured here. The first is the question of how much of the R&D performed in given industries goes toward the production of patented goods in the four technology groups. The second is how much of R&D should be allocated to the German market when innovations are often marketed more broadly. These are dealt with in turn. The assumptions and procedure are similar to that used by Schankerman (1991) in calculating R&D expenditures relevant to the French market.

Total R&D figures for every country (Japan, United States, Western Europe) are for 1975, or the closest year available, and are classified by 23 SIC industry groups. The percentage of R&D by each industry group allocated to each of the four IPC technology groups was based on the assumption that R&D allocation is reflected in patenting. Thus, for each industry, the percentage of R&D going to a given technology was set equal to the percentage of the industry's total patenting going to that technology group. The Yale-Canada patent concordance (B-matrix) was used to find the IPC composition of patenting in the different industry groups, the implicit assumption being that the composition of innovative activities performed within SIC groups is fairly stable across countries. (Note that one would ideally like a concordance for each country.) Thank you to Mr. Jonathan Putnam, Yale University, for extracting and supplying the concordance entries relevant to the four technology groups.

One adjustment to the R&D figures was necessitated by the fact that many pharmaceutical patents have their first IPC classification in chemicals. Based on experience collecting the patent data, it was estimated that approximately three-fourths of pharmaceutical patents have a first classification in pharmaceuticals, with the remaining fourth classified in chemicals. Thus the percentage of total R&D going to the pharmaceutical group was inflated by a third to account for those hidden in chemicals.

Because one important way to exploit innovation is through sales of products incorporating the innovation, the amount of R&D allocated to the German market was assumed to equal the importance of exports to Germany in total sales. (Or, conversely, the importance of the domestic market in total sales for R&D expenditure by German companies.) Sales and export data for each country and the 23 SIC industries are from the IBRD-OECD data base. Mr. George Papaconstantinou of the OECD was very helpful in providing these and the R&D figures.