PATENTS, MARKET STRUCTURE, AND WELFARE.

A Theoretical Investigation into New Dimensions of the Patent System

Thesis submitted for the degree of

Doctor of Philosophy

by

MANFREDI M. A. LA MANNA

LONDON SCHOOL OF ECONOMICS

University of London

ABSTRACT

Chapter I: The relationship between inventor and the Patent Office is modelled as a 'patent regulation game' and it is shown that the conventional wisdom that the P.O. always maximizes welfare by playing the Stackelberg leader is incorrect. Other solution concepts are explored and it is found that, because of the patent life constraint, a reversal of roles may be beneficial. The result that social welfare can be maximized by the P.O. being a Stackelberg follower survives (albeit for a narrower range of values of the key parameters) even if the P.O.-leader is endowed with the additional instrument of a compulsory royalty rate.

Chapter II: A new twist is added to the debate on the Schumpeterian competition hypothesis, by considering the structure of the final-product market as a policy instrument, set by the Patent Office by manipulating patentability standards. It is found that for a vast range of demand functions and under constant returns to scale, a patentability standard that allows for more than one patent to be granted within a given product/process class is welfare superior to the monopoly-generating first-past-the-post current system. If patent life is beyond the P.O.'s control and/or there are increasing returns, no patentability standard is unambiguously preferable.

Chapter III: When Research and Development are modelled as two analytically distinct stages, the choice between patentability standards (whether to grant patents to research prototypes or to fully-developed products) is shown to affect the allocation of resources ***between*** Research and Development. It is shown that under a single-patent regime, granting patents to research prototypes is unambiguously welfare-improving, whereas under a multiple-patent regime a change to patents being granted to fully-developed products and the attending ***increase*** in market uncertainty may ***raise*** welfare.

Chapter IV: The economics of the 'integer constraint' is analysed and it is found that proper treatment of the indivisibility of firms may ***reverse*** the qualitative conclusions of interger-unconstrained models. As an example, a product quality oligopoly model is examined and it is shown that not only the Chamberlinian excess entry result does not apply but also that a free-entry oligopoly and a socially managed industry may produce goods of identical quality, irrespective of the values of cross-derivatives deemed crucial in the literature. Moreover, the integer constraint is shown to provide an explanation for a positive correlation between profitability and concentration in a Cournot oligopoly model ***with free entry***.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

INTRODUCTION

"Is there life after patent life?"

The present dissertation can be read as an attempt to provide an affirmative answer to the above question; more precisely, it is argued that fundamental dimensions of the patent system, other than patent life, are given by **patentability standards**. Albeit in different guises, the first three quarters of the dissertation deal with first defining, and then analysing the welfare effects of, patentability standards.

Chapter I turns on its head the traditional technique for determining optimal patent life, by framing the relationship between the Patent Office and a would-be innovator as a game and analysing the whole range of games generated by different assumptions regarding the strategic roles played by the players (i.e., leader, follower) and their control variables (patent life, patentability standards). It turns out that the solution of the optimal game selection need not coincide with the game traditionally deployed to compute patent life, i.e. the game in which the Patent Office plays the leader and sets patent life. In particular, it is shown that when the process that generates innovations is 'reasonably' productive, social welfare is maximized by letting the innovator determine his own patent life (unsurprisingly, he will set it at an infinite level) and by setting a 'high' minimum patentability standard. The notion of minimum patentability standard is characterized in the context of a new-product model as a shift variable, quantified as the extent by which either the inverse demand curve is pushed upwards or the marginal cost curve pulled downwards.

Chapter II adds a new twist to the debate on the Schumpeterian trade-off, by considering the role of patentability standards in altering the balance between static and dynamic efficiency in a welfare-improving manner. The connection between patents and the Schumpeterian trade-off can be best understood by noting that R&D and the production of output are obviously two different stages and that, *ceteris paribus*, increased competition at each stage is socially beneficial. However, as an increase in competition at one level can only be achieved by reducing competition at the other level, a balance has to be struck between competition and protection from competition. In chapter II the emphasis is on the trade-off implicit in the fact that while an increase in competition at the output stage engineered by granting property rights to more than one innovator brings benefits to the users of the innovation, the resulting non-cooperative behaviour among innovators lowers the expected returns to R&D, thereby curtailing competition at the R&D stage. In the model analysed in this chapter, in which there are many potential innovators, the relevant notion of patentability standard is in terms of **novelty**. Two extreme cases are considered, a **strict** patent regime in which only one patent is awarded within any given product/process class (i.e. the traditional first-past-the-post system), and a **permissive** regime in which all genuine innovators obtain a patent on their discovery, irrespective of priority and novelty considerations (a losers-get-some system). Of course, the latter regime must be cheat-proof, i.e. must not allow patents to be granted to mere imitators, but this turns out not to be an unsurmountable problem. The general conclusion is that neither patent regime is unambiguously superior. Conditions are located under which each regime generates higher welfare levels: if patent life is a policy instrument and there are constant returns to scale in production, then the permissive regime performs better than the traditional winner-takes-all scheme (provided demand is not 'too' convex); whereas the reverse holds if effective patent life is beyond Patent Office's control and the probability of failure at the R&D stage is 'high'.

Chapter III extends the dimensions of the patent system in another direction: by characterizing Research and Development as two analytically **distinct** stages, it poses the question whether patents ought to be awarded to the output of either the research stage (i.e. to research prototypes) or the development stage (i.e. to fully developed products/processes). In terms of patentability standards, the criterion of **industrial applicability** is added to the novelty criterion introduced in the previous chapter. Research is distinguished from Development in terms of both uncertainty of outcomes (the former being stochastic, the latter certain) and of degree of (dis)continuity, in the sense that whereas the outcome of development efforts is assumed to be a continuous, and increasing, function of expenditure in development inputs, a strong threshold effect is assumed in the case of research expenditures, with all investment levels below (above) the threshold yielding no (no increase in the) probability of success in discovery. This simple distinction between R and D is sufficient to endow the patent system with a hitherto unnoticed role - that of conveying valuable information to firms engaged in R&D. In fact, if patents are awarded to research prototypes, investment in development can be undertaken under certainty, whereas if a new product/process has to be fully developed before being patentable, additional uncertainty is generated. Thus the Patent Office, by either reducing or increasing the level of endogenous uncertainty, can alter the allocation of resources **between** research and development. The main result of the chapter is that under a single-patent regime granting patents to research prototypes is always welfare improving, whereas under a multiple-patent scheme, welfare may be increased by generating additional uncertainty through the awarding of patents to fully-developed products/processes.

The final chapter is devoted to a theme that underlies the previous three, namely the importance of the integer constraint brought about by the indivisibility of firms. In the context of a product-quality model it is shown that the integer constraint may reverse the qualitative characteristics of an integer-unconstrained equilibrium. Moreover, it is argued that it can also account for a positive correlation between profitability and concentration in an oligopoly model with free entry.

As the sub-title of the dissertation makes clear, my main aim has been to attempt to extend the economic theory of patents to new dimensions. The motivation behind this choice has not been a lack of interest in the empirical analysis of patents - quite the opposite, in fact. The belief that good empirical work can only be grounded on a foundation of economic analysis that encompasses more than patent life has prompted me to try to enrich the standard model of patents, avoiding, on the other hand, some of the more sterile 'extensions' suggested by certain game theorists, who have used patents and R&D as mere excuses to formulate ingeneous multi-stage games devoid of much empirically relevant economic content.

ACKNOWLEDGEMENTS

A number of people have read and commented at various stages on sections of this dissertation - among them Stanislaw Gomulka, Anthony Horsley, George Norman, and David Ulph. Some of their advice I gratefully took, some I unwisely rejected. The percentage of Type I to Type II errors I leave to the reader to determine. Special thanks are due to David de Meza (with whom I co-authored the article on which parts of Chapter II are based) who has always encouraged me with a tenacity well beyond the duty of a supervisor, and to Max Steuer, whose inflexible notion of a deadline provided the final impetus to achieve a steady state in the production of a definitive version.

Financial support by the Italian National Research Council (CNR), N.A.T.O. (Research Fellowship Scheme), and L.S.E. (Teaching Assistantships Scheme) is gratefully acknowledged.

CHAPTER I

OPTIMAL PATENT LIFE vs OPTIMAL PATENTABILITY STANDARDS

I.1 Introduction

All models of optimal patent life[1] analyse the relationship between the Patent Office and a would-be innovator as a game in which the former sets patent life and behaves as a Stackelberg leader, taking into account the constraint posed by the latter's profit-maximizing behaviour. No justification is provided as to why the above arrangement should be optimal, presumably on the ground that in simple duopoly games leadership is more advantageous than followership and that assigning to the Patent Office the task of setting patent life seems intuitively appealing. The analogy with a duopoly game, though, may be misleading. Unlike a duopoly game, in which both control variables (e.g. prices, output levels) and the players' strategic roles (follower, leader) are exogenously determined, in the context of a patent game it is the Patent Office that sets the rules of the game, i.e. how patents are applied for and granted.

[1] See Arrow (1962), Nordhaus (1969) Ch. 5, Scherer (1972), Stoneman (1987) Ch. 9.

Thus, a fully-optimizing Patent Officer (P.O.) has to solve a double assignment problem, choosing both her role, i.e. either follower or leader, and her control variables, i.e. either minimum patentability standard, σ, or patent life, T. All models of optimal patent life implicitly assume that welfare is always maximized under the game in which the P.O. leads and optimizes over T. However, it seems obvious that the solution of the above game selection exercise will in general depend on demand and technology conditions; and this is, in fact, the conclusion reached in this chapter, where it is shown that when innovations are not 'difficult' leadership is irrelevant and setting patentability standards is more efficient than setting patent life.

I.2 Patent games

Given that the criticism levelled here against 'optimal' patent life models refers to the very way of conceiving the relationship between the P.O. and the innovator rather than to the details of the game, one can successfully appeal to the Occam's razor principle and set out the argument in its simplest terms.

In order to consider the welfare implications of the choice between patent life and patentability standards as policy instruments, a simple model of a new product will be formulated that allows one to quantify the notion of a (minimum) patentability standard.

Let $P(Q)$ be the inverse market demand function for a latent good, i.e a good which under the pre-innovation technology cannot be produced profitably in positive amounts. Let the pre-innovation level of (constant) marginal cost be c and assume that production of Q is marginally unprofitable, i.e. $P(O)+\varepsilon=c$, where ε is 'small'. Then there are two types of innovation that can make the production of a new product profitable:

(i) a 'process innovation' that shifts downwards the marginal cost curve over the relevant output range;

(ii) a 'product innovation' that shift the inverse demand curve to the right over the relevant output range.

We can define the extent by which either $P(Q)$ is shifted outwards or c is reduced as the extent of the innovation, σ, and take it as the patentability standard variable.

The post-innovation equilibrium can be described by the familiar triad of Consumer surplus, Gross profits, and deadweight Loss (as depicted in Fig. I.1), each of which depends on the extent of the innovation σ. The existence of a patent introduces, of course, a time element in the form of the patent term T.

The innovator's objective is to maximize the present value of the stream of profits generated by the innovation, i.e.

$$\pi(\sigma,T) \equiv \int_{o}^{T} e^{-rt}G(\sigma)dt-R(\sigma) \equiv \left(\frac{1-e^{-rT}}{r}\right)G(\sigma)-R(\sigma) \quad \text{I.1}$$

where $G(\sigma)$ is gross profit and $R(\sigma)$ is the Innovation

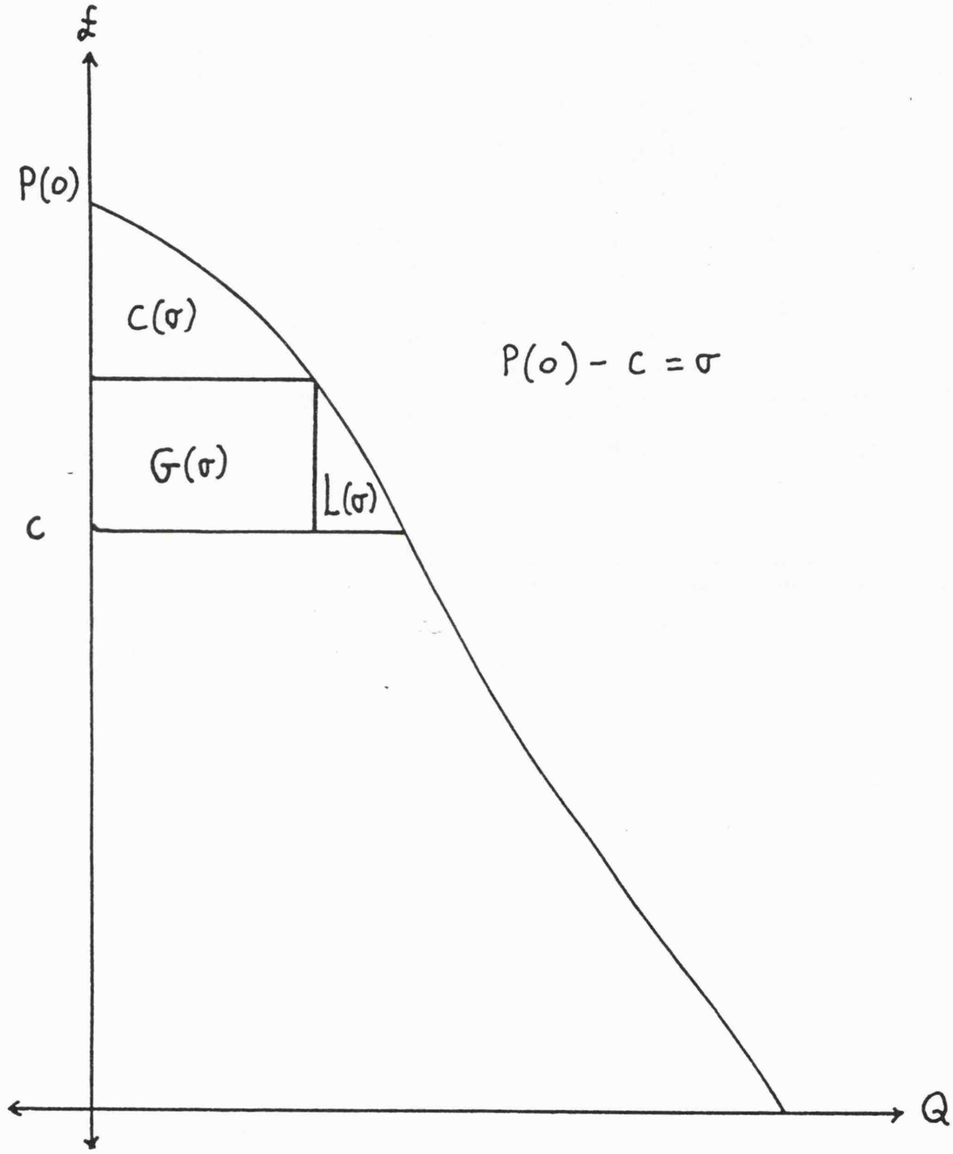


Fig. I.1

Post-innovation Consumers' surplus,

Gross profits, and welfare Loss

Possibility function that maps expenditure in R&D inputs, R, into the extent of the innovation, σ, r is the private (and social) rate of discount and T is patent life.

It is assumed that the Patent Officer is a conscientious civil servant who refrains from pursuing her own welfare and maximizes instead a (well-defined) social welfare function; as the s.w.f. is assumed to be distributionally neutral, the P.O.'s objective is to maximize the present value of the sum of consumers' and producer's surplus:

$$W(\sigma,T) \equiv \int_0^T e^{-rt}[C(\sigma)+G(\sigma)]dt + \int_T^\infty e^{-rt}[C(\sigma)+G(\sigma)+L(\sigma)]dt - R(\sigma)$$

i.e.

$$W(\sigma,T) \equiv [C(\sigma)+G(\sigma)]\frac{1}{r} + L(\sigma)\frac{e^{-rT}}{r} - R(\sigma) \qquad \text{I.2}$$

I.2 implies that after the patent has expired the technology becomes freely available and production of the new good continues under perfectly competitive conditions.

It is more convenient to write $\tau \equiv 1-e^{-rT}$ and, noting that as T ranges from 0 to ∞, τ ranges from 0 to 1, to imagine that there exists a patent life constraint, i.e. $\tau \leq 1$, so that I.1 and I.2 can be written as:

$$\pi(\sigma,\tau) \equiv \frac{\tau}{r} G(\sigma) - R(\sigma) \qquad \text{I.3}$$

$$W(\sigma,\tau) \equiv \frac{1}{r}[C(\sigma)+G(\sigma)] + \frac{1-\tau}{r} L(\sigma) - R(\sigma) \qquad \text{I.4}$$

Having specified the payoffs of the two players, it is revealing to contrast what can defined as the **patent game** played by the P.O. and the innovator with a duopoly game.

Instrument assignment

In any duopoly (indeed, oligopoly) game, the problem of which player sets which variables simply does not arise, for, in a sense, the control variables determine the identity of the decision-maker. For example, in a Cournot oligopoly each firm is defined by its ability to set its own output level. Not so in the patent game, in which the interaction between the two players takes place through two variables (extent of innovation σ, and patent life T), each of which could be controlled by either player, depending on how the rules of the game are set.
The key point here is that, unlike oligopoly games where the definition of the control variables (e.g. either prices or quantities) is exogenous, in the patent game the assignment of control variables is determined by one of the players - by the P.O. herself.

Role assignment

Again in contrast with standard leader-follower oligopoly models in which the role of Stackelberg leader is determined exogenously[2], in the patent game no *ad hoc* assumptions are required to identify the Stackelberg leader, for the decision of which player should play which role is taken by the P.O.

---

[2] The logical difficulties of embedding a leader-follower structure within a static oligopoly model are well known (see Friedman (1983) Ch. 5) but, of course, do not apply to the patent game.

What underlies the difference in instrument and role assignments between the patent game and a duopoly game is that the former takes place in a specific institutional environment designed by one of the players. As soon as it is realized that the P.O. sets the rules under which patents are applied for and granted, it is natural to explore the whole range of instrument and role assignments, so as to check whether the option considered in traditional optimal patent models is the only feasible one and, if not, whether it is superior in terms of social welfare to other feasible games.

The game in which the Innovator plays role R and optimizes over z, with the P.O. playing role S and optimizing over **w** will be referred to as a (Rz, Sw) game. Thus, excluding the double leadership case, there are in principle six games (i.e. six feasible combinations of strategic role and control variable) the P.O. can choose from, as shown in Table I.1.

			PATENT OFFICER			
			Follower		Leader	
			σ	T	σ	T
INVENTOR	Follower	σ	–	Fσ, FT	–	Fσ, LT
INVENTOR	Follower	T	FT, Fσ	–	FT, Lσ	–
INVENTOR	Leader	σ	–	Lσ, FT	–	–
INVENTOR	Leader	T	LT, Fσ	–	–	–

Table I.1

Alternative Patent Games Classified

By Control Variable and Strategic Role

The six patent games of Table I.1 can be grouped into three classes, $\{\langle F\sigma, F\tau\rangle, \langle L\sigma, F\tau\rangle\}$, $\{\langle F\tau, F\sigma\rangle, \langle L\tau, F\sigma\rangle, \langle F\tau, L\sigma\rangle\}$, and $\langle F\sigma, L\tau\rangle$, with each class yielding a distinct equilibrium, as shown below.

Lemma I.1: Irrespectively of the innovator's strategic role, followership by the P.O. combined with optimization over τ yields the economically insignificant (0,0) equilibrium (no innovation, zero patent life).

In both the $\langle F\sigma, FT\rangle$ and $\langle L\sigma, FT\rangle$ games, the first-order condition for the maximization of social welfare is[3]

$$W_\tau(\sigma, \tau) = -L(\sigma)/r < 0 \qquad \text{I.5}$$

and the associated reaction function is

$$\tau^{PO}(\sigma) = 0 \qquad \text{I.6}$$

When the P.O. follows and optimizes over τ, her reaction function is degenerate and coincides with the σ-axis, thereby generating the economically insignificant no-innovation, zero-patent life equilibrium. ■

Lemma I.2: Irrespectively of the innovator's strategic role, optimization over σ by the P.O. yields the $\langle\sigma^*, 1\rangle$ equilibrium, i.e. patent life is infinite ($\tau=1$) and the minimum patentability standard σ^* is determined by $\tau^{PO}(\sigma^*)=1$, where $\tau^{PO}(\sigma)$ is the P.O.'s reaction function.

Consider first the $\langle L\tau, F\sigma\rangle$ game which, it may be noted, is the

[3] In obvious notation, subscripts stand for differentiation.

mirror-image of the traditional 'optimal' patent life model: whereas in the latter the P.O. plays the leader's role, here she choose to follow and, in contrast with the conventional model, allows the innovator to determine patent life, subject to a minimum patentability standard.

The innovator-leader's problem is:

$$\max_{\tau\leq 1} \pi(\tau) \equiv \frac{\tau}{r} G\big(\psi(\tau)\big) - R\big(\psi(\tau)\big) \qquad \text{I.7}$$

where $\psi(\tau) \equiv \tau^{PO^{-1}}(\sigma)$ and $\tau^{PO}(\sigma)$ is the P.O.'s reaction obtained from the first-order condition for the maximization of welfare, i.e.

$$W_{\sigma}(\sigma,\tau) \equiv \frac{1}{r}[C_{\sigma}(\sigma)+G_{\sigma}(\sigma)]+\frac{1-\tau}{r}L_{\sigma}(\sigma)-R(\sigma) = 0 \qquad \text{I.8}$$

$$\tau^{PO}(\sigma) = \frac{1}{L_{\sigma}(\sigma)} [C_{\sigma}(\sigma)+G_{\sigma}(\sigma)+L_{\sigma}(\sigma)-rR(\sigma)] \qquad \text{I.9}$$

It is simple to show that $\pi_{\tau}(\tau)>0$, and thus the patent-life constraint is binding $(\tau=1)$ and the minimum patentability standard σ^{*} is determined by $\tau^{PO}(\sigma^{*})=1$:

$$\pi_{\tau}(\tau) = \frac{G(\sigma)}{r} + \psi_{\tau}(\tau) \ \{\frac{\tau}{r} G_{\sigma}(\sigma) - R_{\sigma}(\sigma)\}>0 \qquad \text{I.10}$$

Inequality I.10 follows from the fact that the quantity in curly brackets equals $-\frac{1}{r}\{C_{\sigma}(\sigma)+(1-\tau)[L_{\sigma}(\sigma)+G_{\sigma}(\sigma)]\}<0$ and $\psi_{\tau}(\tau)$ is the reciprocal of the slope of the P.O.'s reaction function, which is negative in view of the second-order condition for the maximization of welfare.

Under the $(F\tau, F\sigma)$ and $(F\tau, L\sigma)$ games the innovator's first-order condition is

$$\pi_{\tau}(\sigma,\tau) = G(\sigma)/r > 0 \qquad \text{I.11}$$

i.e. the patent-life constraint $\tau \leq 1$ is binding; or, to put it differently, the innovator's reaction function is degenerate (the best-reply level of τ is a constant) and coincides with the $\tau=1$ line.

In conclusion, all three games in which the P.O. optimizes over the minimum patentability standard σ yield the $(\sigma^*, 1)$ equilibrium, characterized by an infinite patent life and a minimum patentability standard σ^* determined by the intersection of the $\tau=1$ line and the P.O.'s reaction function $\tau^{PO}(\sigma)$, shown as point **A** in Fig. I.2. ■

Summarizing the argument so far: we have shown that out of six possible patenting schemes two (namely, $(F\sigma, F\tau)$ and $(L\sigma, F\tau)$) yield the economically insignificant solution of no innovation, three (namely $(F\tau, L\sigma)$, $(F\tau, F\sigma)$, and $(L\tau, F\sigma)$) are feasible and all sustain the same equilibrium, characterized by an infinite patent life and by a minimum patentability standard, σ^*, lying on the P.O.'s reaction function.

The interesting question, of course, is whether this equilibrium can yield a higher level of welfare than the traditional version of the patent game, in which the P.O. acts as a leader and optimizes w.r.t. patent life τ, i.e. the $(F\sigma, L\tau)$ patent game.

In this respect we can prove the following

Theorem I.1: **There is always a range of values of cost and demand parameters such that the $(F\sigma, L\tau)$ patent game yields lower welfare levels than under the alternative game $(L\tau, F\sigma)$.**

Proof. As the second-order conditions for the maximization of profit and welfare (see I.3-4) guarantee that the innovator's and the P.O.'s reaction functions are respectively upward- and downward-sloping in the (τ,σ) plane, to prove the above theorem it suffices to show that they cross at a point where $\tau>1$, or, equivalently, that $\{\tau^{PO}(\sigma^{*})=\tau^{I}(\bar{\sigma})=1\}\Rightarrow(\sigma^{*}>\bar{\sigma})$, as shown in Fig. I.2. In fact, this would imply not only that the alternative patent game $(L\tau,F\sigma)$ performs better than the traditional $(F\sigma,L\tau)$ whenever the latter calls for an infinite patent life, but also (by continuity) that there exists a set of cost and demand parameters such that the two games yield the same level of welfare $W(\hat{\sigma},\hat{\tau}) = W(\sigma^{*},1)$, with $\hat{\tau}<1$, as shown in Fig. I.2. Therefore the $(L\tau,F\sigma)$ game would perform better than the classic $(F\sigma,L\tau)$ game whenever demand and cost conditions were such that the latter game called for an optimal patent life τ^{o}, $\hat{\tau}<\tau^{o}\leqslant 1$.

The relevant first-order condition, and associated reaction function, for the maximization of the innovator's profits are respectively

$$\pi_{\sigma}(\sigma) \equiv \frac{\tau}{r} G_{\sigma}(\sigma) - R_{\sigma}(\sigma) = 0 \qquad \text{I.12}$$

$$\tau^{I}(\sigma) = r\,\frac{R_{\sigma}(\sigma)}{G_{\sigma}(\sigma)} \qquad \text{I.13}$$

Let $\tau^{PO}(\sigma^{*})=1$, i.e., from I.9, $C_{\sigma}(\sigma^{*})+G_{\sigma}(\sigma^{*})-rR_{\sigma}(\sigma^{*})=0$, then

$$\tau^{I}(\sigma^{*}) = \frac{C_{\sigma}(\sigma^{*})+G_{\sigma}(\sigma^{*})}{G_{\sigma}(\sigma^{*})} > 1 \qquad \text{I.14}$$

Define $\bar{\sigma}$ as $\tau^{I}(\bar{\sigma})=1$, then as $\tau^{I}_{\sigma}(\sigma)>0$, $\bar{\sigma} < \sigma^{*}$. ■

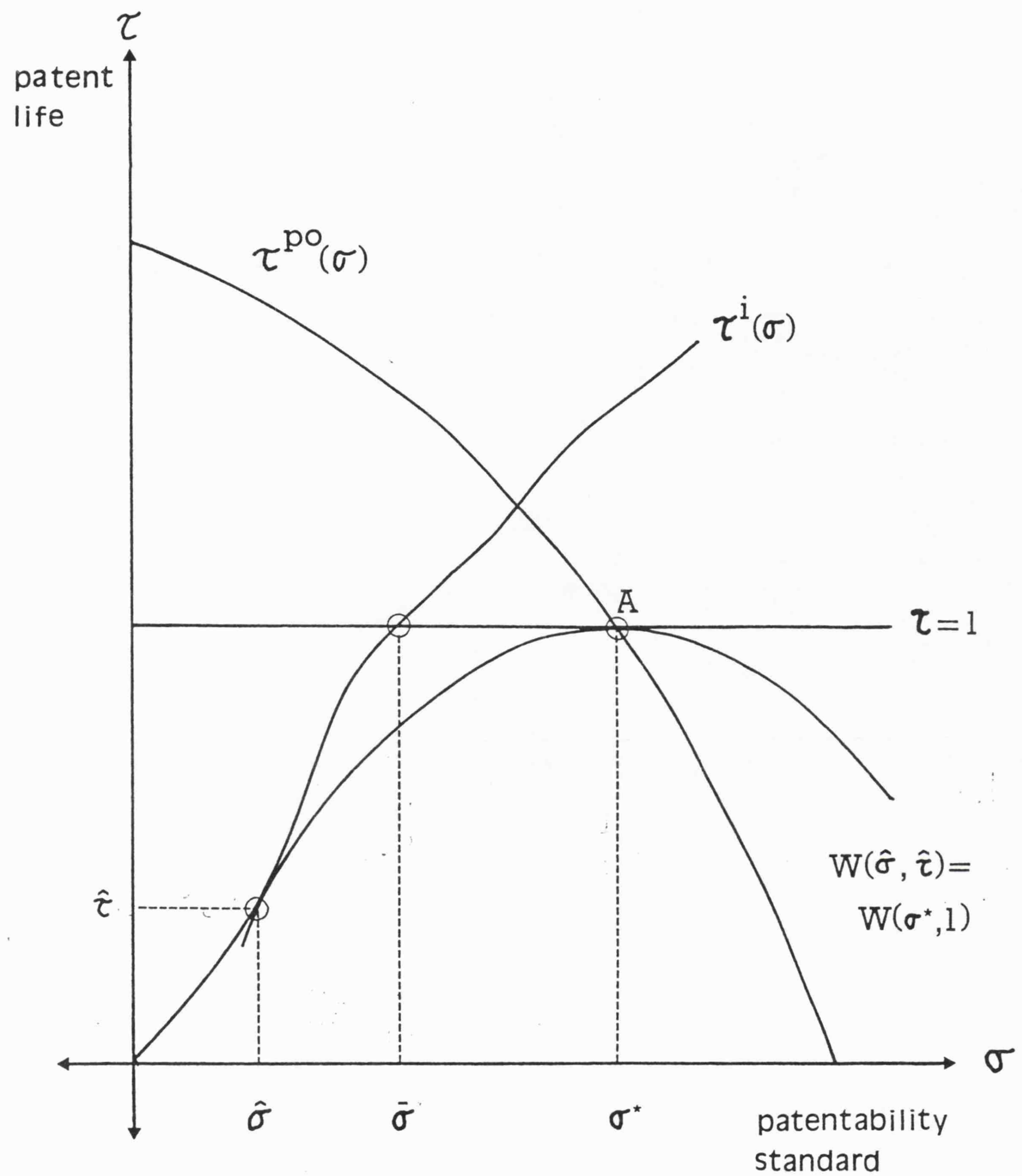


Fig. I.2

Optimal patent life vs optimal patentability standard

I.3 Parametrizations and examples

At this stage it may be useful to resort to some specific cost and demand functional forms so as to ascertain under what conditions which patenting arrangement performs better.
The following parametrizations have been chosen both for their simplicity and because they have been used extensively in the literature and therefore will allow the reader to make comparisons with other models.
The inverse demand function $P(Q)$ is assumed to be linear in output (w.l.o.g. we can take the slope to be unity, for this merely involves a suitable choice of unit of measurement)

$$P(Q) = a - Q \qquad \text{I.15}$$

Both pre-innovation marginal cost is constant and equal to c_o.
The technology for the production of (as yet unspecified) innovation is summarized in an Innovation Possibility Function (IPF), $R(\sigma)$, that maps the extent of the innovation, σ into the expenditure in R&D inputs, R.
The IPF is assumed to belong to the iso-elastic family:

$$R(\sigma) = \theta\sigma^{1/\alpha} \qquad \text{I.16}$$

Depending on whether the innovation is of the 'process' type (and thus the extent of the innovation refers to the shift of the MC curve) or of the 'product' type (which shifts the $P(Q)$ curve) the post-innovation cost and demand curves will be

$$C(Q,\sigma) = c_o - \theta\sigma^{1/\alpha} \qquad \text{I.17}$$

$$P(Q,\sigma) = a + \theta\sigma^{1/\alpha} - Q \qquad \text{I.18}$$

Using I.15-16 and assuming for simplicity that without R&D investment the production of Q is 'marginally' unprofitable (i.e. $a=c_o$), it is elementary to compute Gross profits $G(\sigma)$, Consumer surplus $C(\sigma)$, and deadweight Loss $L(\sigma)$:

$$C(\sigma) = L(\sigma) = 2G(\sigma) = \frac{\sigma^2}{4} \qquad \text{I.19}$$

Therefore the two players' payoffs can be written as

$$W(\sigma,\tau) = \frac{(4-\tau)}{8r}\sigma^2 - \theta\sigma^{1/\alpha} \qquad \text{I.20}$$

$$\pi(\sigma,\tau) = \frac{\tau\sigma^2}{4r} - \theta\sigma^{1/\alpha} \qquad \text{I.21}$$

where, as usual, $\tau \equiv 1-e^{-rT}$.

The two relevant reaction functions required to compare the traditional $(F\sigma, L\tau)$ game and the alternative $(L\tau, F\sigma)$ scheme are as follows:

$$\tau^I(\sigma) = \frac{2\theta r}{\alpha}\,\sigma^{(1-2\alpha)/\alpha} \qquad \text{I.22}$$

$$\tau^{PO}(\sigma) = 4 - \frac{4\theta r}{\alpha}\,\sigma^{(1-2\alpha)/\alpha} \qquad \text{I.23}$$

Under the $(F\sigma, LT)$ regime, in which the P.O. maximizes welfare (I.20) w.r.t. τ, taking $\tau^I(\sigma)$ as a constraint, the (unique) solution is:

$$\tau^* = \begin{cases} \dfrac{8\alpha}{1+4\alpha} & \text{for } \alpha<1/4 \\ 1 & \text{for } 1/4\leq\alpha<1/2 \end{cases} \qquad \sigma^* = \begin{cases} \left(\dfrac{4\alpha^2}{(1+4\alpha)\theta r}\right)^{\alpha/(1-2\alpha)} \\ \left(\dfrac{\alpha}{2\theta r}\right)^{\alpha/(1-2\alpha)} \end{cases} \qquad \text{I.24}$$

Remarks:

1. If innovations are 'easy' (i.e. $1/4 \leq \alpha < 1/2$), the granting of a patent is not an effective way of counter-balancing the output-restricting behaviour of an Innovator-Monopolist. The P.O. hits the $\tau \leq 1$ constraint when the marginal benefit of patent life extension is still positive. Indeed the Innovator is able to attain his first-best optimum and social welfare would be unaffected if patent term were self-administered by innovators.
2. When innovations are 'difficult' ($\alpha < 1/4$) short-lived patents are an effective means to check the propensity to over-invest in R&D by an unregulated Innovator-Monopolist (assuming $\alpha=0.1$, optimal patent life is less than 6 (3) years if r=10% (20%)). Somewhat surprisingly, patents seem to be more efficient in curbing potentially excessive innovation than in promoting it.

To obtain a more complete picture of the circumstances under which the $(L\tau, F\sigma)$ patenting scheme yields an improvement on the conventional $(F\sigma, LT)$ scheme, we may compute the equilibrium value of the extent of innovation under a $(L\tau, F\sigma)$ scheme, σ^*, i.e. $\tau^{PO}(\sigma^*)=1$ for the above parametrization, which turns out to be :

$$\sigma^* = \left(\frac{3\alpha}{4\theta r}\right)^{\alpha/(1-2\alpha)} \qquad \text{I.25}$$

One of advantages of the above formulation is that welfare levels can be fully parametrized in terms of α, the R&D elasticity of the minimum patentability standard, i.e. for any given α we can compute and plot the levels of welfare associated with the equilibrium of each of the two patent games $(F\sigma, L\tau)$ and $(L\tau, F\sigma)$, namely $W(\sigma^{\diamond}, \tau^{\diamond})$ and $W(\sigma^{*}, 1)$; see Fig. I.3.

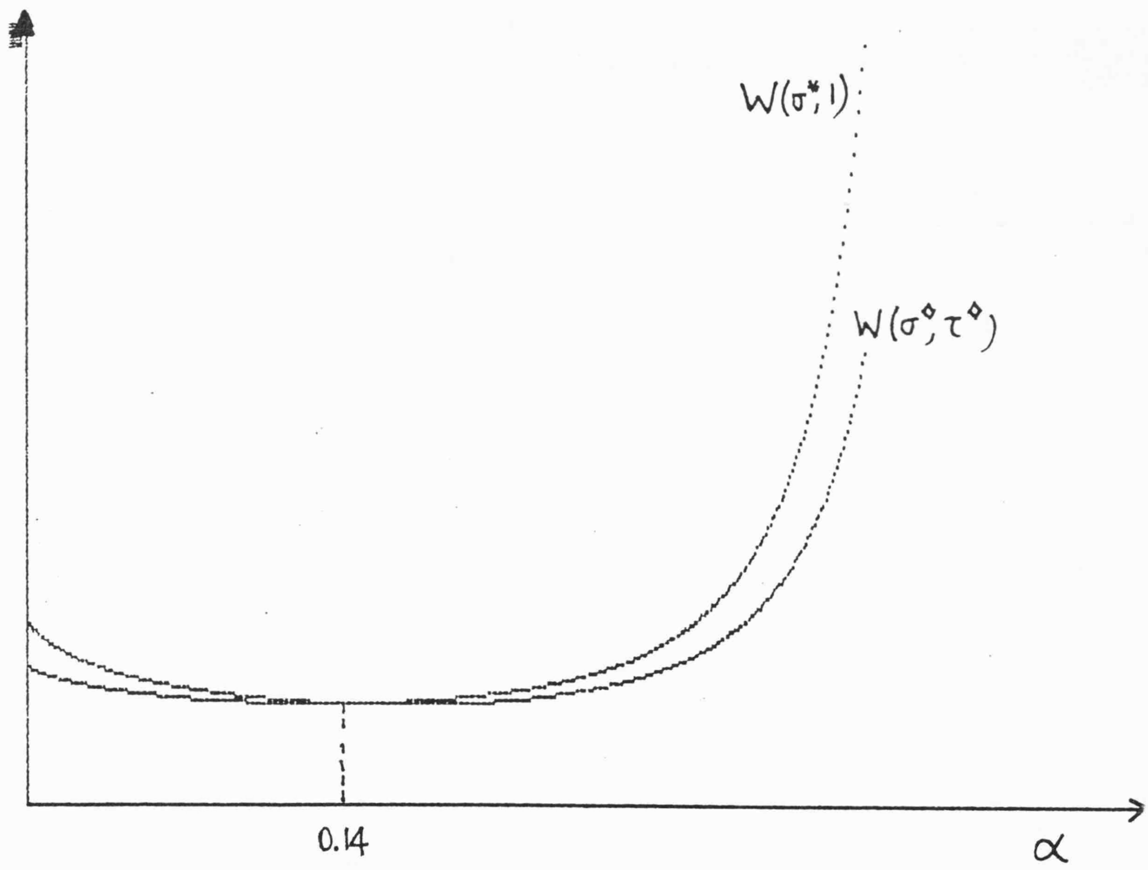


Fig. I.3

The $(L\tau, F\sigma)$ game vs the $(F\sigma, L\tau)$ game

Fig. I.3 reveals that the alternative game yields an improvement on the conventional arrangement, not only when innovations are easy (i.e. $1/4 \leq \alpha < 1/2$) and thus patent life is infinite under both regimes, but also when innovations are 'not too difficult' (i.e. $1/7 < \alpha < 1/4$).

In conclusion, this simple parametrization has the advantage of showing rather dramatically that the presence of a binding constraint on the control variable set by the leader may nullify and even reverse the benefit of being a leader at all. The reason for this rather intriguing fact can be better understood if player j's reaction function $\tau^{J}(\sigma)$ is interpreted as an **innovation schedule** that specifies the minimum patentability standard σ required to be granted a patent term τ. When innovation are not 'difficult', the social benefit flowing from the ability to set a high minimum patentability standard σ^{*} more than offsets the cost of letting the innovator choose an infinite patent life.

I.4 Comparisons with Tandon's model

Another, albeit indirect, way of testing the robustness of the claim that the $(L\tau, F\sigma)$ game may yield a substantial improvement on the $(F\sigma, L\tau)$ game is to compare the former with an 'augmented' version of the latter, in which the P.O. plays the leader and sets not only patent life, τ, but also a compulsory royalty rate ρ. Such an 'augmented' model has been analysed by Tandon (1982), who, using a model very similar to the one sketched above, obtains two main results:

1. optimal patent life is always infinite;
2. the optimal royalty rate, ρ^*, is the solution of the following equation:

$$\frac{4\alpha-1}{\alpha}\rho^2 - 3\rho + 1 = 0 \text{ [4]} \qquad \text{I.26}$$

It might have been thought that for a model in which the P.O. is endowed with **two** policy instruments (τ and ρ) the implicit assumption of conventional models that the $(F\sigma, L\tau)$ game is always the most beneficial could indeed be valid.

On the contrary, using the above parametrization, it is simple to show that although endowing the P.O. with a second instrument does narrow the range of vales of α such the the alternative $(L\tau, F\sigma)$ scheme is welfare superior, it does not eliminate it altogether.

Letting ρ be the compulsory royalty rate expressed as a percentage of the minimum patentability standard σ, social welfare and the innovator's profits can be written respectively as

$$W(\sigma,\tau,\rho) = \frac{(1-\rho^2\tau)}{2r}\sigma^2 - \theta\sigma^{1/\alpha} \qquad \text{I.27}$$

$$\pi(\sigma,\tau,\rho) = \frac{(1-\rho)}{r}\rho\tau\sigma^2 - \theta\sigma^{1/\alpha} \qquad \text{I.28}$$

---

[4] In Tandon's article, the optimal ρ is defined by the following cubic (where $x \equiv (1-\alpha)/\alpha$):

$-(x-3)\rho^3 + (x-6)\rho^2 + 4\rho = 0$; however, by factoring out the economically meaningless solution $\rho=1$, we obtain I.26.

Note that for $\rho=½$ the Tandon scheme boils down to the traditional $(F\sigma, L\tau)$ game.

Denoting the equilibrium values under the Tandon scheme and under the $(L\tau, F\sigma)$ game with a bar and with a star respectively, we obtain

$\bar{\tau} = \tau^* = 1$ (infinite patent life)

$$\bar{\rho} = ½\{\frac{3\alpha}{4\alpha-1} + [(\frac{3\alpha}{4\alpha-1})^2 - \frac{4\alpha}{4\alpha-1}]^{½}\} \qquad \text{I.29}$$

$$\bar{\sigma} = [\frac{2\alpha}{r\theta}\,\bar{\rho}(1-\bar{\rho})]^{\alpha/(1-2\alpha)}; \quad \sigma^* = (\frac{3\alpha}{4\theta r})^{\alpha/(1-2\alpha)} \qquad \text{I.30}$$

Fig. I.4 plots the values of $W(\bar{\sigma}, \bar{\tau}, \bar{\rho})$ and $W(\sigma^*, 1)$ for all admissible values of the parameter α and shows that for 'easy' innovations (i.e. for $0.28 \leq \alpha < ½$) the P.O.'s advantage of being endowed with an additional policy instrument is more than offset by the disadvantage of not being able to set a high minimum patentability standard.[5]

---

[5] As is the case for most models of regulated monopoly, an obvious criticism to any model of optimal patent life is that, if the regulator (i.e. the P.O.) has all the information required to compute the 'optimal' value of her instruments (i.e., patent life and/or compulsory royalty rate), she will not choose to regulate the monopolist in the prescribed manner, for she would be able to attain a first-best solution by command. In the context of the model sketched above, this means that the P.O. would merely instruct the Innovator to invest the socially optimal amount in R&D, reward him with a lump-sum and then market the new good at marginal cost. Tandon's dismissal of the above criticism on the ground that lump-sum rewards are "totally impractical" is ill-founded, as his own model does not provide any explanation as to why lump-sum compensation should be unfeasible. If lump-sums are impractical (and they are) it must be on account of factors ignored by Tandon (and by most of the literature). The inherent contradiction of

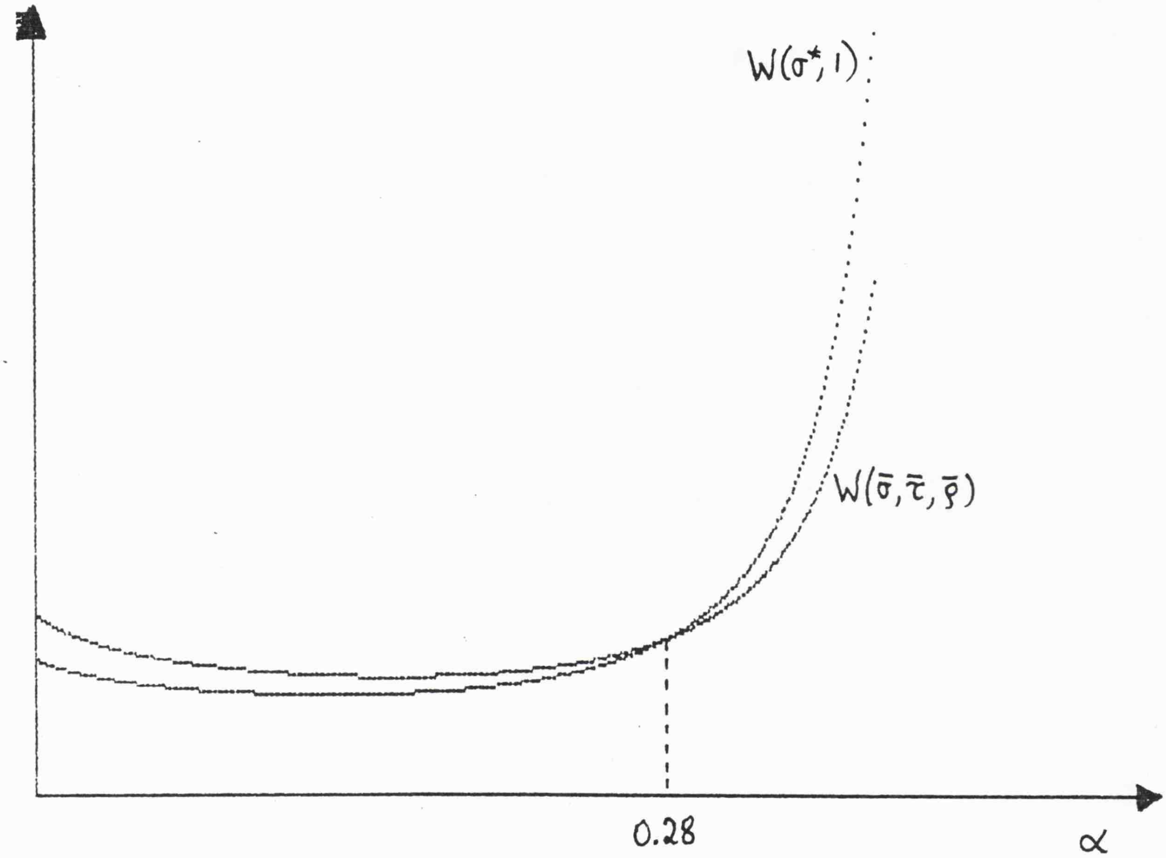


Fig. I.4

The $(L\tau, F\sigma)$ game and Tandon's scheme

---

optimal patent life models is that on the one hand as long as they are cast in a complete-information framework they cannot provide a rationale for regulating innovators by patent-related means (e.g. patent life, compulsory royalty rates, etc) and on the other, as soon as incomplete information is introduced they fail to provide an optimal regulatory mechanism.

CHAPTER II

Winner takes all vs losers get some: the Schumpeterian trade-off and patentability standards

II.1 Introduction

Among the supporters of the patent system as a socially beneficial institution, few have been more eulogistic than Bentham :

> [A patent] ... *unites every property which can be wished for in a reward. It is variable, equable, commensurable, characteristic, exemplary, frugal, promotive of perserverance, subservient to compensation, popular and revocable.*[1]

Even Bentham was aware that the "winner-takes-all" feature of the patent system could result in a wasteful duplication of inventive efforts, but believed that the net effect was still in favour of a "first-past-the-post" reward structure for inventions. It is interesting to note that Bentham lists three rather feeble arguments in support of his verdict in favour of a winner-take-all patent system:

---

[1] J. Bentham, *The Rationale of Reward*, London, 1825, p. 92.

(i) the pain of disappointment after trial is more than offset by the pleasure of expectations before it;

(ii) in any given field of inventions there are not "many" potential competitors;

(iii) losers may reap indirect benefits ("develop their talent").

Point (i) is based on the questionable assumption that interpersonal comparisons of cardinal utilities can be scientifically validated; point (ii) is an unwarranted assumption, especially in view of the widespread phenomenon of 'cluster technologies' and of the fact that the number of potential competitors can be altered by manipulating the rules whereby patents are awarded (i.e. the fields of inventions are not given); finally, point (iii) misses the crucial fact that the indirect benefits to the losers can be profitably turned into direct benefits by changing the very structure of the patent system.

The aim of this Chapter is to show that there exist changes in the patent system that, by increasing the pay-offs to the losers, can combine all the desirable properties described by Bentham with a reduction in the net welfare losses created by patents.

This Chapter is concerned with an old dilemma:

Suppose that a monopolist, faced with the prospect of being granted a patent that generates a net profit of £π, would invest $£x_1$ in R&D and charge a price of p_1 per unit of patented good. Consider now the alternative arrangement whereby each of n firms is allowed to patent the same good: each firm would earn less than $£\pi_1/n$ (unless collusion is perfect), would invest $£x_n$ ($<£x_1$) in R&D and charge p_n ($<p_1$). Which market structure is socially preferable?

A huge literature - admirably surveyed by Kamien and Schwartz (1982) and Baldwin and Scott (1987) - has attempted to model and investigate empirically the trade-off underlying the above dilemma. The so-called Schumpeterian trade-off, of course, is between "dynamic efficiency" (related to the allocation of resources to R&D) and "static efficiency" (related to the allocation of resources, for given levels of R&D): a monopoly generates a higher rate of technical change but also larger welfare losses due to higher prices. It would take an uncommonly large dose of self-esteem to believe that anyone can add anything of even marginal interest and originality to the vast literature on Schumpeterian competition (Baldwin and Scott (1987) cite 340 references). The surprising fact, that I believe warrants a further addition to the literature, is that the role of the patent system in **altering** the above trade-off seems to have escaped notice or even mention.

The issues at stake can be put in a better focus if it is recognized that R&D activity and manufacture of final goods are obviously two different stages of the production process and that *ceteris paribus* increased competition at each stage is socially beneficial. What makes the Schumpeterian trade-off interesting is that increased competition at one level is incompatible with increased competition at the other so that the problem is not one of raising the degree of competitiveness, but rather that of striking the right balance between competition and protection from competition. In this perspective the system whereby protection from competition is established (i.e. the patent system) can be seen as a key element of the Schumpeterian trade-off. In this Chapter I shall analyse the trade-off implicit in the fact that, while raising the degree of competitiveness at the output stage by granting property rights to more than one inventor does benefit the users of the invention(s), the resulting non-cooperative behaviour, by lowering the expected returns to R&D, will discourage entry into the race, thus curtailing competition at the R&D stage.

Thus it is a genuine policy dilemma whether it is preferable to restrict competition at the output stage by granting exclusive property rights in invention to only one firm thereby encouraging entry (and competition) at the R&D stage or to award more than one patent, reaping the benefits of increased competition on the final good market and foregoing the benefits of a larger pool of potential inventors.

The Chapter is organized as follows: in the next section I shall consider the practical and theoretical problems of designing an optimal 'novelty' patentability standard; in sec. II.3 a simple model of R&D and patenting is formalized; sec. II.4 reports the main results; sec. II.5 highlights the relevance of the preceding analysis for patent reform, especially in the field of biotechnology; finally, Appendices II.A-E contain either special cases or complementary issues of the general analysis.

II.2 The 'novelty' patentability criterion[2]

The patently obvious fact that defining a 'novelty' patentability criterion is indispensible for the very existence of the patent system ought not to deserve mention, were it not that only very recently economists have begun to address the question of designing an optimal novelty standard. In fact, it is clear that the whole patent system would be completely subverted if the novelty requirement were enforced so as to grant a patent to any new variant, however trivial, of a genuine invention; at the other extreme, a strict interpretation of the novelty criterion could paralyse technological progress (e.g. television as 'radio with pictures'). The reason why economists have been reluctant to face the problem of defining 'novelty' is, of course, the difficulty of characterizing the 'similarity' between

---

[2] A condensed version of sections II.2-4 has appeared in La Manna *et al.* (1989).

goods: in the Arrow-Debreu world there is no way of determining in a meaningful way the degree of 'closeness' between size 9 shoes, size 9½ shoes, and chocolate bars.[3]

It is not surprising, if not excusable, that in the once-recurring debates on the patent system, the alternative had been between the complete abolition of the patent system and its retention with an unspecified novelty criterion.[4]

There are two modelling strategies to analyse the problem of assessing the effects of alternative novelty criteria. As an example of the first approach consider the following case: an inventor, by expending R&D resources, produces a single good located at point **a** on the product line for which aggregate demand in the absence of any other variety would be $D(p)$. Consumers are located on the product line with a given distribution of the 'transport cost' of substituting alternative varieties for the patented good and with a given distribution of reservation prices. In this context, defining a novelty criterion means determining the size of the patent-holder market, i.e. the 'width' **w** of the patent, in the sense that no competing firm can produce a variety within a distance **w** of the patented good. The best example of this approach can be found in Klemperer (1989), who has characterized the optimal duration/breadth mix

---

[3] For an interesting attempt to define a topology appropriate to characterize product quality, see Horsley (1982).

[4] For an insightful account of the early controversies on patent law, see Machlup and Penrose (1950).

in the context of a model of horizontal product differentiation (HPD) and has established that, under fairly robust conditions, the optimal mix depends on the relationship between the distribution of reservation prices across consumers and the distribution of their transport (or utility mismatch) costs. The HPD approach to defining the optimal novelty standard emphasises the additional welfare loss due to a wide patent breadth, namely the loss suffered by consumers whose preferred varieties fall within the specification of the patented good and who have to settle for the cheaper unpatented good. The HPD approach, though, is unsuited to analyse how the incentive to undertake costly R&D investment can be altered so as to determine the optimal extent of patent protection in the case of a number of essentially identical inventions. The idea of competition underlying the HPD approach is that for any potential product/process class there is at most one firm engaged in R&D, whose profits have to be protected from the erosion due to cheaper copies being made available to consumers by imitators who do not contribute anything to technical progress. The notion of competition underlying the alternative approach taken in this Chapter is quite different in so far as it is assumed that at any one time there is a number of firms engaged in R&D activity so that the problem for the Patent Office is not that of protecting a single inventor from plagiarists, but rather to regulate competition among genuine inventors. A good example of an industry whose characteristics are closer to our approach than to the product differentiation approach is the biotechnology industry, in which firms have a clear idea of the desired invention and thus direct their R&D effort towards a fairly narrow segment of the relevant product space:

" A British court has just decreed that Genentech, a biotechnology company in California, cannot retain exclusive marketing rights in Britain for its heart product, TPA. The judge ruled that the terms of the patent were too broad (...) To stop others working (...) would stifle research and not be in the public interest. Genentech plans to appeal against the decision. If its patent were to stick, 19 companies would have to abandon their work on their TPAs."

[*The Economist*, July 18, 1987]

Most of the theoretical literature on patents has little to contribute towards addressing the question posed in the above quotation, in so far as it has hitherto failed to consider any intermediate solution between a strict first-past-the-post system and the complete abolition of patents. The problem is that, even though in practice even late finishers generally obtain some positive payoff to their R&D effort, formal models of patent races typically assume that the winner takes all.[5]

---

[5] Well-known examples are Dasgupta and Stiglitz (1980), Loury (1979), Lee and Wilde (1980).

In this Chapter I shall examine whether it is socially desirable to adopt policies which increase the extent to which the rewards to R&D are shared between the participating firms. Before examining the welfare implications of a reformed patent system, it has to be established whether it be at all feasible to widen the distribution of returns to R&D without abolishing altogether the patent system. The problem, of course, is that if patents do not protect genuine inventors from mere plagiarists (i.e firms that refrain from investing in R&D and simply imitate the improved process/product), the incentive to undertake R&D activity would be dramatically reduced with potentially damaging welfare effects.[6] A simple way of excluding free-riders and allowing genuine inventors who fail to arrive first to benefit from their R&D activity is the following 'open registry scheme'. Suppose that a patent is awarded **T'** periods after filing (under

---

[6] Suppose, at one extreme, that the patent system were abolished. Incentives to undertake R&D would still remain, as argued long ago by Plant (1934). Production lags, enhanced by industrial secrecy, give early inventors temporary market power before copies appear. Indeed, it is well documented that researchers sometimes fail to apply for patents on their discoveries even though they are available [see, e.g. the references in Levin (1986)]. Now, in the absence of patents, a genuine inventor who independently replicates a discovery, say a month after the first to find it, does benefit because even with two producers super-normal profits will generally be available. This feature is in contrast to the reward structure implicit in the traditional patent system in which the independent inventor who completes the research project a little late receives nothing (at least if patent life is of any length) just as does the plagiarist who simply seeks to copy the work of others.

the current patent system the administrative lag between filing and being awarded a patent averages 2½ years); any application for a product/process within the same class is granted a patent provided it is filed for before the award of the patent to the first claimant, with the obvious provision that details of patent applications are not disclosed before the date of the award. As in many cases production of the patented good lags several years from the granting of the patent,[7] the open registry scheme need not delay the production of the new good/ process.

Is there any merit in a losers-take-some reward structure ? Given risk aversion it follows that, other things equal, sharing the rewards will enhance the attractiveness of R&D and so increase the flow of R&D. As demonstrated in Appendix II.A this may well be an important benefit, but is not the effect explored here, for risk neutrality is assumed. Even so, the risk reduction implicit in a reward structure that shares the returns may have real effects. With a conventional patent system, all that matters is getting home first and so risk-neutral firms engaged in a patent race have an incentive to gamble on a risky research strategy [see Klette and de Meza (1986a)]. Completing

---

[7] In the next chapter it will be argued that in many cases welfare can be increased by stipulating that patents be awarded to research prototypes requiring substantial post-patent development investment, thereby making any delay between filing for patents and production of no consequence.

the course in the expected time virtually guarantees that the race will be lost, at least if there are many competitors. When all that matters is winning, it is better to select a bold strategy that yields the possibility of a very fast time even if there is a even greater chance of complete failure. It makes no matter whether you are second or last. In a R&D context, the consequence is that R&D strategies turn out to be excessively risky from a social viewpoint. Although inventing a week before a rival has a high private payoff, the social advantage is slight. Introducing multiple prizes diminishes the cost of not being first and therefore leads to a socially preferred choice of research strategy. However, this potential benefit is also excluded from the present analysis, for firms are assumed to be unable to influence the riskiness of their R&D plans.

The concern of this Chapter is with the trade-off implicit in the fact that, while allowing more than one inventor to profit from R&D leads to non-cooperative behaviour which diminishes the expected return to R&D, given the volume of R&D, competition among inventors will benefit the users of the invention. The key issue is whether it is better to encourage competition in the discovery or the dissemination of improved products/processes. As suggested above, this is a genuine dilemma and thus it is not surprising that it cannot be determined unambiguously whether introducing multiple prizes is desirable or not. Indeed, it is slightly surprising that fairly general conditions have been found under which rewarding late finishers is welfare improving.

The model developed here is highly stylized. The purpose is to suggest that the winner-takes-all feature of the conventional patent system is not necessarily the best possible arrangement. In practice the existing patent system often does give some reward to late finishers and thus lies between the two polar cases we analyse. The purpose is to identify whether it is desirable to introduce this feature deliberately. The results are likely to be robust, but it must be recognized that many practical details will be treated cursorily. Nevertheless, this Chapter makes the case that it is worth taking alternatives to the present patent system seriously.

II.3 The model

The general structure of the model follows that of Dasgupta and Stiglitz (1980). Risk-neutral firms spend an amount **x** on R&D at time 0 and this yields a known probability of inventing at each subsequent moment.[8] For most of their analysis Dasgupta and Stiglitz specify a Poisson distribution of invention times, but

---

[8] The assumption of a fixed-size R&D project is undemanding: as shown by Dasgupta & Stiglitz, even if **x** is variable, its equilibrium value is always the same (provided there are many entrants).

for my purpose I can afford not to specialise the density function. Firms follow independent R&D strategies and seek a particular new product. At moments at which there is but a single producer it earns monopoly profits of π_1. When there are $n>1$ producers collusion is imperfect and each earns gross profits of π_n, with $n\pi_n<\pi_1$. As we assume that R&D is aimed at producing a new product (i.e. a product which prior to invention cannot be sold at a profitable price) and that all potential producers are equally efficient, it is easy to see that a successful multiplant inventor will never choose to licence other firms to produce the new product.

Two patent regimes will be considered. Under the **strict** regime the first to invent receives a patent lasting for T_s years. During this period no other firm receives a patent and so cannot produce. When the patent expires there is freedom of entry into the industry. Under the alternative **permissive** regime, patents are awarded to all genuine inventors (i.e. non-plagiarists) as long as they apply within T' years of the first claimant. It can be seen that this formulation embodies the open-registry scheme discussed in the introduction. One obvious problem with the permissive scheme is that it pays for any one inventor to buy up all other patents thereby earning monopoly profits. However, this drawback is unlikely to carry much force, not only because, as shown by Salant *et al.* (1983), the strategy of buying up competitors is profitable only if a very large proportion of patents is secured, but also because the antitrust authorities would have little difficulty in enforcing a restriction that a firm cannot own more than one patent on the same product/process. It can then be safely assumed that monopolisation is not feasible.

A permissive regime with patent life T_p allows for multiple producers even during the life of the first patent granted. Once the first patent expires the idea is public property and, because of the long lead time, new production can start quickly, rendering all other patents economically worthless.

Let there be **n** entrants to the R&D stage under the strict regime and write the density function that the first discovery occurs t_1 periods after time 0, the second after t_2 periods and so on as $P(t_1,..,t_n)$. Under a strict regime (using a formulation that appears cumbersome but is subsequently useful) a free-entry equilibrium satisfies :

$$E\{P(t_1,..,t_n)\pi_1 g(t_i)\} - nx = 0 \qquad \text{II.1}$$

where

$$g(t_i) = \begin{cases} 1/r[e^{-rt_i} - e^{-r(t_i+T_s)}] & i=1 \\ 0 & \forall i>1 \end{cases} \qquad \text{II.2}$$

r is the discount rate and π_1 is instantaneous monopoly profit.

Consider now the permissive regime. Since there is normally a positive probability that the first inventor will face competition within the life of its patent and assuming that in that event collusion is imperfect, it follows that with no change in patent life or the number of entrants, expected profits will be lower than under a permissive regime. Therefore, at a free-entry equilibrium the permissive regime, *ceteris paribus*, sustains fewer firms, **m**<**n**. Instantaneous gross per-firm profits when **m** firms are active are π_m and $m\pi_m < (m-1)\pi_{m-1}$. If **m** firms enter the R&D race, then under the permissive regime

$$E\{P(t_1,\ldots,t_m)\sum_{i=1}^{m} i\pi_i\ f(t_i)\} - mx = 0 \qquad \text{II.3}$$

where

$$f(t_i)=\begin{cases} 1/r[e^{-rt_i}-e^{-rt_{i+1}}] & t_1+T'>t_{i+1} \quad i\geq 1 \\ 1/r[e^{-rt_i}-e^{-r(t_1+T_p)}] & t_{i+1}>t_1+T'>t_i \quad i\geq 1 \\ 0 & t_i>t_1+T' \quad i>1 \end{cases} \qquad \text{II.4}$$

II.4 The Case for Permissive Patents

The two patent regimes can now be appraised according to their net expected social welfare, defined as the sum of expected consumers' surplus and industry profits, net of R&D costs:

Proposition II.1 : **The permissive regime is welfare superior to the strict regime if patent life is a policy variable, there are constant returns to scale in the production of the new good, and market demand is not 'too convex' (e.g. linear or of constant elasticity).**

Proof : The general form of the proof is to show that if T_s is the optimal patent life in the strict regime then in a permissive regime there exists a patent life $T_p > T_s$ which attracts the same number of entrants to the race and yields greater social benefits. The advantage of stipulating that the number of entrants be the same under the two regimes is that attention can be restricted to social welfare gross of R&D expenditures. As, *ceteris paribus*, the free-entry number of firms under the permissive regime is less than under the strict regime, it follows that, in order for the two scheme to sustain the same number of firms, **n**, patent life under the permissive regime must be increased, say to T_p. It is not obvious that there exists a T_p sufficiently large to induce the required amount of additional entry, but in Appendix II.B it is shown that this is not a problem. Here it is simply assumed that a T_p with the required property does exist.

From II.1 and II.3 we obtain

$$E\{P(t_1,\ldots,t_n)\{[f(t_1)-g(t_1)]+\sum_{i=2}^{n}f(t_i)i\pi_i/\pi_1\}\}=0 \qquad \text{II.5}$$

Now consider the welfare implications of a switch from the strict to the permissive regime. Aggregate social efficiency, net of R&D costs, at each moment is the sum of consumers' surplus plus industry profits. This is clearly maximized when price equals marginal cost. Actual welfare may therefore be written as the difference between this maximum level of welfare, W^*, and D_i, the temporary deadweight loss incurred for the duration of the patent when there are **i** firms active in the market. It follows that

$$E\{W^{s}\}\equiv E\{P(t_1,\ldots,t_n)[(W^{*}/r)e^{-rt_1}-D\,g(t_1)]\}-nx \qquad \text{II.6}$$

$$E\{W^{p}\}\equiv E\{P(t_1,\ldots,t_n)[(W^{*}/r)e^{-rt_1}-\sum_{i=1}^{n}D_i f(t_i)]\}-nx \qquad \text{II.7}$$

These equations embody the fact that when the first patent expires deadweight loss is zero if at least one firm has invented. From II.6-7 it can be seen that a switch from the strict to the permissive regime yields a welfare gain if :

$$E\{P(t_1,\ldots,t_n)\{[f(t_1)-g(t_1)] + \sum_{i=2}^{n}f(t_i)D_i/D_1\}\} > 0 \qquad \text{II.8}$$

From II.5 and II.8 it is evident that the switch raises welfare if

$$i\pi_i/\pi_1 > D_i/D_1 \qquad \forall i\leq n \qquad \text{II.9}$$

Although we have not characterized optimal patent life under a permissive regime, we have identified the inequality that has to be satisfied in order for a switch to a permissive regime to be welfare-improving. To ascertain under what conditions inequality II.9 holds, some specializiation of the demand functions is required, because knowledge of local properties is insufficient to evaluate welfare losses. Appendix II.C shows that II.9 holds when demand is linear in output or of constant elasticity. That II.9 holds provided demand is not 'too convex' can be shown as follows.

In Appendix II.D it is easily established that under Cournot assumptions

$$\frac{d\log n\pi_n}{dn} > \frac{d\log D_n}{dn} \quad \text{iff} \quad 1-n+\frac{[P(n)-c]Q(n)}{D_n}>0 \qquad \text{II.10}$$

where $P(n)$ and $Q(n)$ are respectively price and output in a n-firm Cournot oligopoly.

When $n=1$ the above inequality is always satisfied and as II.10 is simply the continuous version of II.9, this suggests that if the rules of the permissive regime restrict the number of patentees to two at most, then it must be preferable to a strict regime. However, as the change from $n=1$ to $n=2$ is not small, II.10 could turn out to be a poor approximation of II.9. However, it can be shown that in order for Proposition II.1 not to hold in the continuous case, demand has to be 'very convex'. Notice that, as the deadweight loss 'triangle' D is less than the rectangular area $\Delta Q(P-c)$, where ΔQ is the increase in market demand when price falls from the n-firm oligopoly level $P(n)$ to the competitive level c, inequality II.10 can be written as

$$1-n+\lambda Q/\Delta Q > 0 \qquad \text{II.11}$$

where $\lambda \equiv (P-c)\Delta Q/D > 1$.

Under Cournot assumptions $P(n)-c$ equals $P(n)/\varepsilon n$, where ε is the point elasticity of demand evaluated at $P(n)$.

Defining the arc-elasticity of demand ε^* as $(\Delta Q/\Delta P)(P(n)/Q(n))$, where $\Delta P=P(n)-c$, it follows that II.11 implies

$$\frac{\varepsilon}{\varepsilon^*} > \frac{n-1}{\lambda n} \qquad \text{II.12}$$

So, if II.10 is to fail

$$1 > \frac{n-1}{\lambda n} > \frac{\varepsilon}{\varepsilon^*} \qquad \text{II.13}$$

i.e. $\varepsilon^*>\varepsilon$; this is turn requires that over the relevant range the elasticity ε be increasing in Q. This imposes a severe restriction on the curvature of the inverse demand curve P(Q); in fact it rules out concave, linear and constant-elasticity demand curves and requires P(Q) to be strongly convex. In fact, straightforward differentiation reveals that in order for $d\varepsilon/dP$ to be negative, the elasticity of the slope of the inverse demand curve, i.e. $P''(Q)/\big(P'(Q)Q\big)$ must exceed $(1+\varepsilon)/\varepsilon$, i.e. P(Q) has to be strongly convex.

If P(Q) is strongly convex, the deadweight loss D is significantly less than $\frac{1}{2}(P-c)\Delta Q$. Hence λ is certainly greater than 2.

To violate II.11 it is thus required that $\varepsilon^* \geq \frac{2n}{n-1}\,\varepsilon$.

This is the basis for claiming that **unless the point elasticity of demand rises very rapidly as price falls, a permissive regime is preferable to a strict patent regime.** ■

It is interesting to note that Proposition II.1 holds even if the instantaneous probability of success depends only on current (per-period) R&D investment, rather than, as in the Dasgupta & Stiglitz model on R&D costs sunk at the start of the programme.

A possibility not explored here as a third alternative between a strict and a permissive regime is the case of a strict patent regime coupled with compulsory licensing. It could be argued that the latter, by avoiding the R&D costs incurred by late finishers while the registry is open and by imposing a maximum price, can combine the advantages of the two 'pure' patent regimes. However, it must be recognized that a licensing scheme suffers from a number of problems that do not occur under a permissive scheme, i.e.:

(i) by sharing rewards, the permissive patent arrangement cuts the real cost of risk, if, as is almost certainly the case, inventors are risk averse. This benefit would be lost under a compulsory licence scheme which still preserves the winner-takes-all feature;

(ii) a compulsory licence scheme gives an inventor an incentive not to apply for a patetn at all and rely instead on secrecy. Resources will be expended to maintain secrecy and, in any case, the period of protection afforded byt secrecy is unlikely to be ideal from a social point of view;

(iii) especially if the patent is granted to a research prototype (see next Chapter), a compulsory licence that discloses the barest details of a new process/product may be of limited value if the inventor is not willing to cooperate by providing ancillary information;

(iv) complusory licences will entail resource costs in monitoring revenues, outputs, etc. to pay the fees accurately;

(v) as mentioned in sec. II.1, late finishers will not normally come up with identical ideas, even though under the existing rules these may not be sufficiently novel to be granted a patent. Still, the additional variety may well be enough to justify the R&D costs expended during the possibly short period of time the registry remains open.

Thus, the superiority of a strict-cum-licenses scheme over the permissive regime is a matter of conjecture, to be decided on empirical grounds.

In Proposition II.1 it is stipulated that there be constant returns to scale in production (i.e. excluding the fixed cost of R&D); the reason for this is that by allowing for economies of scale the balance between to two regimes is tipped against the permissive regime, owing to factors that are independent of R&D itself. It is trivial to show that:

Proposition II.2; A strict regime may be superior to a permissive regime if economies of scale in production are sufficiently great.

Now put everything in reverse. Let T_p^* be the optimal patent life under the permissive regime. Choose T_s^* to yield the same number of firms. Following the same procedure that leads to II.9, the strict regime is certainly welfare superior if

$$i\pi_i^*/\pi_1^* < D_i^*/D_1^* \qquad \forall\ n \geq i \geq 2 \qquad \text{II.14}$$

where D_i^* is the difference between net social surplus when there is a free-entry equilibrium in the producing industry (i.e. the equilibrium when knowledge is public) and realized aggregate benefit when there are only i firms producing. Notice that, because of economies of scale, it is possible that a free-entry equilibrium sustains too many firms and hence $D_i^*<0$.
Appendix II.E shows that with linear demand curves and a fixed production cost, inequality II.14 may be satisfied. The economic intuition underlying Proposition II.2 is simply that in the presence of economies of scale, production efficiency benefits from concentration and this may more than offset the losses from reduced price competition.

* * *

The argument deployed to show under what conditions a switch from the strict to the permissive regime may generate welfare gains breaks down is patent life is not a policy instrument, as it would be the case whenever the economic life of patents is shorter than their legal span. In fact, succeeding waves of technical change may render existing discoveries economically obsolete; as a result, many firms taking out patents have little expectation that the invention will be of commercial value for the full legal life of the patent. Under these circumstances it would be impossible to set $T_P>T_S$, in so far as patent life would have to be the same under the two regimes and beyond the

Patent Office's control.[9]

There is another reason for interest in the case of equal patent life. The abolition of the patent system altogether would result in an environment similar to that of the permissive regime albeit with, in effect, a short T_p. If it can be shown that the welfare effects of a switch in regimes are ambiguous when $T_p=T_s$, this will remain true for some $T_p<T_s$. It can now be shown that:

Proposition II.3: **If patent life is the same under a permissive and a strict regime then, even with constant returns in production, it is ambiguous which is socially preferable.**[10]

In order to prove Proposition II.3 it is sufficient to use a simpler version of the model sketched above. Upon payment of a research fee, **x**, a firm is given a lottery ticket that yields at some future moment a particular product-innovation with probability **p**. If the firm does not invent at that time, it never will. Under the strict regime the (single) patentee reaps the reward flowing from his invention in the form of gross monopoly profits, π_1. In the event of **m** firms 'striking gold', each is awarded π_1 with probability **1/m**. If there are **N** entrants and entry is free, the following zero expected profit

---

[9] Of course a change in regime may affect the date at which the next generation of products is expected to appear, but this complication will be ignored.

[10] A less general proof can be found in La Manna *et al.* (1989), pp. 1435-7.

condition will hold :

$$E\{\pi^s\} \equiv \sum_{i=0}^{N-1}\binom{N-1}{i}p^{i+1}q^{N-(i+1)}\frac{\pi_1\tau}{(i+1)}x=0; \quad p+q\equiv1 \qquad \text{II.15}$$

i.e.

$$E\{\pi^s\} \equiv \frac{1-q^N}{N}\,\pi_1\,\tau - x = 0 \qquad \text{II.16}$$

where τ $\left(\equiv(1-e^{-rT})/r\right)$ is the discount factor and T is patent life.

Under the permissive regime, as all successful inventors are granted a patent, gross profits will depend on the number of successful inventions. Once again it is assumed that collusion is not feasible and that, if there are multiple patentees, they will engage in a Cournot game. With free entry and n entrants, the zero expected profits condition ensures that :

$$E\{\pi^P\} \equiv \sum_{i=0}^{n-1}\binom{n-1}{i}p^{i+1}q^{n-(i+1)}\pi_{i+1}\tau - x = 0 \qquad \text{II.17}$$

i.e.

$$E\{\pi^P\} \equiv \sum_{i=1}^{n}\binom{n}{i}p^i q^{n-i}\, i\pi_i\tau - x = 0 \qquad \text{II.18}$$

where π_{i+1} is gross per-firm profits in a i-firm oligopoly. Letting $\xi_i\equiv\tau\left(i\pi_i+L_i\right)$, where L_i is the aggregate deadweight welfare loss in an i-firm oligopoly, and w^* the welfare gain after the patent has expired and price has been driven to marginal cost, then at a free-entry equilibrium the ratio of net expected welfare under the two regimes can be written as:

$$\psi(p)\equiv\frac{E\{W^s\}}{E\{W^P\}} = \frac{(1-q^N)\left[w^*/r - \xi_1\right]}{(1-q^n)w^*/r - \sum_{i=1}^{n}\binom{n}{i}p^i q^{n-i}\xi_i} \qquad \text{II.19}$$

Proposition II.3 follows from the two following Lemmata:

Lemma II.1: Under certainty the permissive regime is welfare superior.

$$\psi(1) = \frac{w^*/r - \xi_1}{w^*/r - \xi_n} < 1 \qquad \text{II.20}$$

Notice that under certainty the permissive regime not only generates smaller welfare losses (as $\xi_1 > \xi_n$) but also lower aggregate research expenditures (as under free entry, of course, $Nx > nx$).

Lemma II.2: As p approaches 0 the strict regime becomes welfare superior.

Applying de l' Hospital's rule it is easily established that

$$\lim_{p \to 0} \psi(p) = \frac{N}{n} > 1. \ \blacksquare \qquad \text{II.21}$$

The economic rationale for Lemma II.2 is straightforward: when inventions are 'difficult' (i.e. p is 'small') the benefit flowing from the larger entry generated by the strict regime in the form of an increased probability of having at least one successful invention may more than offset the double disadvantage of the strict regime, i.e. the higher research expenditure and the lower gross social surplus if an invention does occur.

II.5 Conclusion

The debate over the social net benefits of the patent system, in focusing on the stark alternative between a no-patent system (which allows perfect free-riding) and a 'strict' patent regime (which prevents genuine but late inventors from benefitting from their own R&D investment), has ignored the mid-way option of a 'permissive' regime. Such a system excludes true free-riders (i.e. those who have not invested in R&D) but does not penalize genuine inventors for not arriving first.

The permissive regime could be instituted by changing the rules by which patents are awarded. Taking into account that there exists an administrative lag between filing for a patent and obtaining it, a simple way of implementing a permissive regime would entail accepting all applications up to the date of the award of a patent to the earliest inventor of a given class of new products/processes. Of course, it must be assumed that, as under the current (strict) regime, the Patent Office will not divulge the technical details of patents before they are awarded even if under the permissive regime there would exist an incentive to bribe patent officers.

A no-patent regime can be interpreted as a permissive regime with patent lives shorter than production and imitation lags. Instead of disclosing early in return for a patent, inventors would minimize the flow of pre-production information and potential free-riders would have to wait until the product appear before copying it. This would give the initial inventors an interval in which to reap supra-normal profits and, more importantly, returns would not be limited to the first past the post. The length of time that inventors enjoyed protection from free-riding would probably be shorter than at present, but since it is already quite common for firms not to avail themselves of patents, this would not necessarily be a crippling blow. Of course, there are some products for which a patent remains valuable for the whole of its legal life. In such cases, the effective shortening of protection afforded to the inventor would be important and may more than offset the advantages of sharing the returns. However, it is worth noting that inventions that are of durable economic value are likely to be major inventions in the sense of generating large cost falls or significantly improved products. It is precisely such inventions that Klette and de Meza (1986b) show should anyway be rewarded with the shortest patents.

In the past, serious economic cases have been made that patents should be abolished because they over-reward the inventor [e.g. Plant (1934)]. What does not seem to have been appreciated is that abolishing patents tends to increase the rewards to coming second or third.

Although no unequivocal case can be made, I hope to have offered a persuasive argument that, for reasons that have previously been neglected, a 'permissive' patent regime, or even a no-patent regime, deserves serious consideration for, under reasonable assumptions, it may be socially preferable to what has traditionally been seen as the ideal of a 'strict' regime.

The above analysis can provide a useful starting point to assess the suitability of existing patent law for the growth and success of one of the most promising areas of industrial research - biotechnology. The problem can be put in focus by considering two examples, the patentability of sexually propagated plants and patent protection for micro-organisms.

Economists trained to treat quality as a scalar or, at most, as a two-dimensional vector may well find bewildering the fact that under European patent law for a variety of barley to be considered "new", only one out of its fifty "descriptors" has to be different from existing varieties; change, by backcrossing, the length of leaf hairs, and you have a "new" variety. In the terminology of this Chapter, the application of the novelty criterion to plant varieties is definitely "permissive". Given the (very) high probability of success in what amounts to "cosmetic breeding", one may surmise that Lemma II.1 may apply.

Similar problems in defining a "novel" product are encountered in the case of micro-organisms; in the well-known case of alpha-interferon, which was first patented by Biogen (and Hoffmann-La Roche) in 1980 in Europe and a few months later in the U.S. by Genentech, the latter's gene differed from Biogen's for a string of twentyfour aminoacids out of hundreds. It seems almost certain that Genentech did not copy Biogen's patent; in industrial biotechnology competing companies have very clearly defined research targets and thus it is not wholly surprising that two very similar patents were filed within a short period of time. In this case, it would appear that the open-registry scheme mentioned above could provide a workable and more desirable alternative to the current system, which combines the worse of both world. In fact, under the European Patent Convention, companies may delay publication of their patent application for eighteen months after the date of priority application. This means that if the application is successful and is given wide "coverage" resources expended by competitors on R&D on similar projects during the 18-month "black-out" period would have been wasted and the patent would have failed to provide a valuable "early warning signal" (see next Chapter).

Appendix II.A

In this Appendix it is shown that under risk aversion a mild relaxation of the strict patent regime which allows the first two inventors to be awarded a patent (rather than only one) may yield a net welfare gain. In order to focus on the specific effects of risk aversion, all other potential benefits flowing from a permissive patent regime are ignored; in particular it is assumed that if two or more firms succeed at the R&D stage the two eventual patentees share monopoly profits π, thereby eliminating the benefits flowing from competition at the output stage. Upon payment of an R&D fee, φ, each entrant acquires a lottery ticket yielding a given new product with probability p. Let $u(y-\varphi)$ be each entrant's utility function, where y is the gross payoff of the game, which, of course, depends on the nature of the patent regime: under a strict, first-past-the-post regime $y=\pi$, whereas under a more permissive two-patent regime if two firms succeed, each receives $y=\pi/2$. The utility function is normalized so that $u(0)=1$.

Under a strict patent regime, assuming that if k firms succeed at the R&D stage each has a probability $1/k$ of reaping π, the equilibrium number of entrants, n, will satisfy the following condition:

$$u(\pi-\varphi)\ \frac{1-q^n}{n} = 1; \quad p+q\equiv 1 \qquad \text{IIA.1}$$

Under a permissive two-patent regime, the equilibrium number of entrants **N** is determined by the following:

$$pq^{N-1}u(\pi-\varphi)+u(\frac{\pi}{2}-\varphi)p\sum_{i=1}^{N-1}\binom{N-1}{i}p^{i}q^{N-(i+1)}\frac{2}{i+1} = 1 \qquad \text{IIA.2}$$

where $\binom{N-1}{i}p^{i+1}q^{N-(i+1)}$ is the probability that a successful firm will be faced with other i successful inventors and the probability of being awarded one of the two available patents equals $\frac{2}{i+1}$. Notice that

$$\sum_{i=1}^{N-1}\binom{N-1}{i}p^{i+1}q^{N-(i+1)}\frac{2}{i+1} \equiv \frac{2(1-q^{N})}{N} - 2pq^{N-1}. \qquad \text{IIA.3}$$

If the equilibrium under a permissive two-patent regime sustains more firms than under a strict regime, then there will be a net welfare gain from the higher probability of a discovery being made at all. The following simple example shows that under risk aversion **N** may indeed exceed **n**.

Suppose that **n**=2; then from IIA.1 and IIA.2, a two-patent regime will yield a 50% increase in the number of entrants (i.e. N=3) provided that

$$pq^{2}u(\pi-\varphi)+[\frac{2}{3}(1-q^{3})-2pq^{2}]u(\frac{\pi}{2}-\varphi) \geq (\frac{1-q^{2}}{2})u(\pi-\varphi). \qquad \text{IIA.4}$$

It is simple to confirm that IIA.4 holds iff[11]

$$\frac{u(\frac{\pi}{2}-\varphi)}{u(\pi-\varphi)} \geq \frac{3}{4}. \qquad \text{IIA.5}$$

For appropriate value of π and φ IIA.5 will be satisfied if the income elasticity of the utility function is less than **1/3**.

---

[11] Notice that IIA.4 holds for $\forall p>0$; however, this result does not extend to all **n** and **N**.

Although the less risk-averse inventors are, the less dramatic the effect on entry, the permissive regime always encourages entry.

APPENDIX II.B

In the proof of Proposition II.1 it is assumed to be possible to find a T_p which induces as much entry as does T_s. This requires that T_s is finite. It is easily shown that it normally will be. The demonstration that follows is for a zero discount rate, however this is no problem because it can be shown that T_s is decreasing in r.[12]

Let the expected date of first discovery be $t=t(n)$. Granted that freedom of entry always results in zero expected profits for the firms, the social problem is to minimize the net loss to consumers from delaying the cheap availability of the invention. The problem is thus to minimize

$$S = tS_o + T_sS_1 \qquad \text{IIB.1}$$

where S_o is the gain in consumer surplus when the new good is introduced at a price equal to marginal cost, as opposed to S_1 which is the gain in consumer surplus when price falls from the monopoly level to marginal cost. Minimizing S requires

$$\frac{\partial S}{\partial T_s} = S_1 + S_o \frac{dt}{dn}\frac{dn}{dT_s} = 0 \qquad \text{IIB.2}$$

Now, entry is determined by the zero profit condition :

$$\pi_1 \frac{T_s}{n} - x = 0 \qquad \text{IIB.3}$$

---

[12] There is an error in Dasgupta and Stiglitz's (1980) formulation of the optimal patent life problem. This is corrected in Klette and de Meza (1986a) who show T_s to be a decreasing function of r.

Equation IIB.3 implies that for some finite patent lives R&D costs are sufficiently low to induce multiple entry. However the Dasgupta and Stiglitz approach makes no sense unless this is true.

Differentiating IIB.3

$$\frac{dn}{dT_s} = \frac{n}{T_s} = \frac{\pi_1}{x} \qquad \text{IIB.4}$$

From IIB.2 and IIB.4, at an optimum

$$S_1 + S_0 \frac{dt}{dn} \frac{\pi_1}{x} = 0 \qquad \text{IIB.5}$$

However, when T_s is large and hence so is **n**, **dt/dn** will be a negative number of small absolute value – when there are already many firms seeking to invent entry of another one cannot advance the expected dateof first discovery by much. Thus for **n** 'large' the LHS of IIB.5 must be positive. For IIB.5 to hold patent life must be lowered thereby raising **dt/dn** and reducing the aggregate deadweight loss. Optimal patent life is finite under a strict regime.

Granted that T_s is not infinite, it is still not certain that even an infinite T_p will induce the entry of **n** firms under the permissive regime. But instead of allowing an unlimited number of patents, if the reward was limited to, say, the first two inventors to succeed, it is virtually certain that if T_s induces **n** firms to enter then, except in the extreme case of Bertrand competition and perfect substitutes, duopoly profits would not be so low as to preclude **n** firms with a sufficiently large T_p.

This is of relevance because Proposition II.1 would hold even if permissive is interpreted as two firms rather than one.

Appendix II.C

The iso-elastic market demand case

In obvious notation, inverse market demand is written as

$$p = \left(\sum_{i=1}^{n} q_i\right)^{-1/\varepsilon} \qquad \text{IIC.1}$$

Let W_m be social welfare, net of production costs (but gross of research costs), when m oligopolists are active, i.e.

$$W_m \equiv \frac{\varepsilon}{\varepsilon-1} \left(\sum_{i=1}^{m} \hat{q}_i\right)^{(\varepsilon-1)/\varepsilon} - c\sum_{i=1}^{m}\hat{q}_i \qquad \text{IIC.2}$$

where $\hat{q}_i$ is the profit-maximizing level of output of firm i in an m-firm Cournot oligopoly and c is marginal and average cost, i.e.

$$\pi_i(\hat{q}_i) \equiv \max_{q_i} \left\{q_i\left[q_i+\sum_{i\neq j}^{m} q_j\right]^{-1/\varepsilon} - cq_i\right\} \qquad \text{IIC.3}$$

It is simple to confirm that at a symmetric equilibrium per-firm output and per-firm profits are given respectively by

$\hat{q}=[m\varepsilon c/(m\varepsilon-1)]^{-\varepsilon}/m$ and $\pi_m(\hat{q})=[m\varepsilon c/(m\varepsilon-1)]^{1-\varepsilon}/(\varepsilon m^2)$ IIC.4

Define W^* as the first-best level of gross social welfare, i.e.

$$W^* \equiv \max_{Q} \left\{ \frac{\varepsilon-1}{\varepsilon} Q^{(\varepsilon-1)/\varepsilon} - cQ\right\} = \frac{c^{1-\varepsilon}}{\varepsilon-1} \qquad \text{IIC.5}$$

Finally, define D_m as the deadweight loss associated with an m-firm oligopoly as compared with the first-best level of social welfare :

$$D_m \equiv W^* - W_m = \frac{1}{\varepsilon-1} c^{1-\varepsilon}\left\{1-\left((m\varepsilon+\varepsilon-1)/m\varepsilon\right)\left((m\varepsilon-1)/m\varepsilon\right)^{\varepsilon-1}\right\} \qquad \text{IIC.6}$$

For Proposition II.1 to hold, it has to be shown that

$$m\pi_m/\pi_1 > D_m/D_1 \qquad \text{IIC.7}$$

i.e., using IIC.4 and IIC.6

$$\frac{\left[(m\varepsilon-1)/m\varepsilon\right]^{\varepsilon-1}/m}{\left[(\varepsilon-1)/\varepsilon\right]^{\varepsilon-1}} > \frac{\{1-\left[(m\varepsilon-1)/m\varepsilon^{\varepsilon-1}\right.\}/m-\left[(m\varepsilon-1)/m\varepsilon\right]^{\varepsilon}}{1-\left[(\varepsilon-1)/\varepsilon\right]^{\varepsilon-1}-\left[(\varepsilon-1)/\varepsilon\right]^{\varepsilon}} \qquad \text{IIC.8}$$

which holds for $\forall m \geq 2$. ∎

The linear demand case

Using the same notation as above it is easy to confirm that if the inverse market demand is linear, i.e.

$$p = a - \sum_{i=1}^{m} q_i \qquad \text{IIC.9}$$

then Proposition II.1 holds because

$$\frac{m\pi_m}{\pi_1} = \frac{4m}{(m+1)^2} > \frac{4}{(m+1)^2} = \frac{D_m}{D_1} \qquad \text{IIC.10}$$

APPENDIX II.D

Write industry profit as

$$n\pi_n \equiv \pi(n) \equiv \big(P(n)-c\big)Q(n) \qquad \text{IID.1}$$

where $Q(n)$ is industry output. Hence

$$\frac{d\log\pi(n)}{dn} = \left(1 + \frac{dP}{dQ}\frac{Q}{P-c}\right)\frac{dQ}{dn}\frac{1}{Q} \qquad \text{IID.2}$$

$$D \equiv \int_c^P Q(z)\,dz - (P-c)Q \qquad \text{IID.3}$$

and thus

$$\frac{dD}{dP} = -(P-c)\frac{dQ}{dP} \qquad \text{IID.4}$$

$$\frac{d\log D}{dn} = -\frac{(P-c)}{D}Q\frac{dP}{dn}\frac{1}{P}\frac{dQ}{dP}\frac{P}{Q} \qquad \text{IID.5}$$

But as

$$\frac{d\log P}{dn} = \frac{dP}{dQ}\frac{Q}{P}\frac{dQ}{dn}\frac{1}{Q} \qquad \text{IID.6}$$

$$\therefore \frac{d\log\pi(n)}{dn} > \frac{d\log D}{dn} \quad \text{iff} \quad \left(1+\frac{dP}{dQ}\frac{Q}{P-c}+\frac{(P-c)}{D}Q\right)>0 \qquad \text{IID.7}$$

Equation II.10 of the text follows from IID.7, recalling that at a Cournot equilibrium

$$-\frac{dP}{dQ}\frac{Q}{P-c} = n$$

APPENDIX II.E

An example that satisfies inequality II.14

Suppose that T_P is such that only two firms enter the race (n=2) and that a free-entry equilibrium would sustain three firms if R&D were free. It is readily calculated that in the linear demand case fixed costs must amount to π_1^e, where π_1^e are monopoly profits gross of fixed costs. It is easy to verify that net monopoly and duopoly profits are respectively $\pi_1 = \frac{3}{4}\pi_1^e$ and $\pi_2 = \frac{7}{36}\pi_1^e$. The deadweight losses associated with π_1 and π_2 (taking as a bench-mark the free-entry equilibrium) are given by

$$D_1^* = -\frac{3}{8}\pi_1^e \; ; \quad D_2^* = -\frac{17}{72}\pi_1^e$$

It follows that

$$\frac{2\pi_2}{\pi_1} = \frac{14}{27} < \frac{D_2^*}{D_1^*} = \frac{17}{27}$$

as required for II.14 to hold.

CHAPTER III

Patents as Early Warning Devices in a Two-Stage Model of Research and Development

III.1 Introduction

In this Chapter I shall formulate a model of R&D in which Research and Development are not seen, as in most of the literature, as analytically identical, but rather as two distinct stages. Having thus provided a richer (and more realistic) characterization of the innovation process, I shall consider whether the Patent Office could make use of the distinction between research and development by manipulating patentability standards in a welfare-improving manner. In particular, I shall address the question of whether patents should be granted to the outcome of either the research stage (e.g. to research prototypes) or the development stage (e.g. to fully-developed products/processes). In order to provide a full welfare ranking of patent regimes, in answering the above question I shall analyse both the options considered in the previous chapter, i.e. I shall examine both single- and multiple-patent schemes.

The general conclusion of the analysis is that, if research and development are two separate stages and patents are granted to the research prototypes, then patents can take on a new role - that of conveying valuable information to firms. The information patents convey does not pertain to technical knowledge but to market structure: by granting patents to research prototypes the Patent Office can remove uncertainty on the structure of the final-product market. Thus, by either reducing or increasing endogenous market uncertainty, the Patent Office can alter the allocation of resources **between** research and development; it is then shown that whereas under a single-patent regime the reduction in uncertainty engendered by granting patents to research prototypes is (almost) always welfare improving, under a multiple-patent regime, expected net social welfare can be increased by granting patents to fully-developed products/ processes, i.e. by introducing **additional uncertainty** into the system.

The Chapter is organized as follows: in sec. III.2 the lamentably short literature on two-stage R&D models is surveyed; in sec. III.3 a two-stage model of research and development is formulated; sec. III.4 offers a full taxonomy of patent regimes according to two patentability criteria ('industrial application' and 'novelty'); sec. III.5 analyses the welfare implications of granting patents to either research prototypes or fully developed products under a single-patent regime; sec. III.6 considers the effects on industrial structure and social welfare of a multi-patent regime in which patents are granted to research prototypes, which are then contrasted in sec. III. 7 with the implications of granting multiple patents to fully developed products/processes. Finally, in sec. III.8 it will be suggested that the above analysis could be profitably applied to the formulation of patent reform in areas such as biotechnology. All technical details are relegated to footnotes.

III.2 'R&D': Unravelling the 'R' from the 'D'.

In the field of the economic theory of R&D few authors have been as candid as Brander and Spencer (1983) in acknowledging that "there is nothing in [the] model that formally distinguishes between cost-reducing R&D and investment in capital stock" (p. 226): this admission points rather starkly to the fact that most R&D models use research and development at best as an example (and at worst as an excuse) to analyse the effects of 'lumpy' investment, disregarding the specific features of the innovative process. It is therefore not wholly surprising to note that within the large set of 'R&D' models, the subset of models that distinguish in a meaningful way between research and development is very small and the subset of 'proper' R&D models that envisage a significant role for patents comes dangerously close to being a singleton, as witnessed by the following survey of the relevant literature.

Nelson (1982). An interesting attempt to relate R&D capabilities to a well-defined notion of 'knowledge' can be found in Nelson (1982), who addresses an even more fundamental issue than the distinction between research and development, by providing various examples of how 'knowledge' can be quantified as a focusing device that allows R&D to be conducted more efficiently. However, although close in spirit to the model formulated in the next section, Nelson's model does not deal with the interaction between research, development, appropriability (as determined, among other things, by patentability standards), and inter-firm rivalry that forms instead the core of our model.

Reinganum (1985). The article by Jennifer Reinganum on 'A Two-Stage Model of Research and Development' provides a good example of the inventiveness of game theorists in using R&D as a vehicle to devise sophisticated multi-stage games with counter-intuitive implications.

Resorting to the analogy of patent races with track races, it can be said that Reinganum's is a sprint race with a peculiar rule for the start-off: before being allowed to sprint, all racers have to run a one-hurdle course and, provided (at least) one participant clears the hurdle, then all participants can line up for the sprint with no handicaps - the only advantage for the winner of the hurdle race being that of having the option of holding the starter's gun and thus being able 'to move first'.

The surprising result is that the winner of the hurdle race (i.e. in the research stage) will choose not to move first and will run at a lower pace in the sprint race (i.e. will invest less in development).

The price Reinganum pays to make the R&D process fit into her ingenious game is to rule out any informational role for patents in affecting the allocation of resources between research and development by assuming that "research findings rapidly become common knowledge" (p. 276): this, of course, presumes that research findings cannot be patented. Although by defining research as sufficiently 'basic', one can always attach to research output the property postulated by Reinganum, at the other end of the R&D process, it is almost never the case that development expenditures cease with the granting of a patent, as assumed in her model. Thus the (tenuous) distinction drawn by Reinganum between research and development cannot address the question of how developed a product/ process should be in order to be patentable, e.g. whether patents should be granted to research prototypes or fully developed products.

Grossman and Shapiro (1987). The first (and to the present writer's best knowledge, only) model that both provides a characterization of research and development as two meaningfully distinct stages and considers the role of patent policy is to be found in Grossman and Shapiro (1987). Their model can be summarized as follows: two firms are engaged in a race to discover a new product/process and **two** breakthroughs are required to win the race, the per-unit-of-time cost of achieving a breakthrough with probability p being c(p). Grossman and Shapiro call research (development) the stage leading to the first (second) breakthrough; the first breakthrough has no intrinsic value, but has to be attained in order to progress to the next stage. Although one might question the assumption of research and development being equally uncertain and costly (i.e. c(p) is the same in both stages), it can be accepted as a first approximation. The crucial assumption made by Grossman and Shapiro relates to the feasibility of research monitoring: unlike a track race, in which monitoring one's rivals' position is feasible and (almost) costless[1], in a multi-stage research race firms cannot observe each others' research efforts (almost by definition of independent research). Thus, Grossman and Shapiro's assumption that "each firm can observe the state of progress of its rival, i.e. whether the rival has successfully completed the first stage of research or not" (p. 374) is highly questionable. Not only it is difficult to envisage situations in which firms know the state of progress of their rivals but not how the progress

---

[1] In races other than sprints, an early leader has the disadvantage of being constrained (by virtue of the location of the eyes at the front of the head) in his or her ability to monitor constantly his or her rivals' positions.

has been achieved, but, more significantly, in Grossman and Shapiro's model each firm has a positive incentive to misrepresent its current state of progress. As a result, their analysis of the implications of granting a patent to the 'intermediate result' (i.e. the outcome of the first research breakthrough) misses one fundamental role of patents, namely to convey information on the firms' state of progress of research. Moreover, in considering the welfare implications of granting patents to either the intermediate result or the finished product, a key factor is bound to be the effect on entry to the race: again, Grossman and Shapiro's exogenously given duopoly structure robs the patent system of other important function, namely to encourage (or discourage) entry.

Horstmann, MacDonald, and Slivinski (1985). Although not directly relevant to the two-stage model of the next section, mention should be made of perhaps the only model to treat patents as information transfer mechanisms, namely the article of the same title by Horstmann, MacDonald and Slivinski (1985) in which it is assumed that success at the R&D stage (they do not distinguish between the two) gives the winner private information on the profitability of the various options open to a competitor (i.e., exit, imitation, and duplication). As a result, the decision whether or not to patent is taken on strategic grounds, i.e. taking into account that the very act of patenting allows the rival to revise his or her expectations regarding the profitability of the various R&D options. Thus, the information being transferred by patents does not pertain to market structure (as in the model analysed in Secs. III.3-7).

Green and Scotchmer (1989). Finally, a recent paper by Green and Scotchmer ought to be mentioned, for it introduces, albeit in an altogether different model from the one discussed below, the notion of an optimal 'patentability standard'. The context is a two-period, two-firm, sequential innovation model. The per-period value (willingness to pay) of period-one innovation is $\mathbf{x}$, has a distribution $F(\mathbf{x})$, and a fixed cost c_1. Patenting the period-one innovation allows an improved version to be made by period-two innovator with a cost of c_2, value $\mathbf{y+x}$, with $\mathbf{y}$ being distributed according to $G(\mathbf{y})$. Normalizing marginal cost to zero and considering the case of Bertrand competition, let $\bar{y}$ be the minimum patentable improvement on period-one innovation, i.e. any improvement $y<\bar{y}$ cannot be marketed without prior permission by period-one innovator. The optimal patentability standard, $\bar{y}^*$, guarantees that the period-two innovator's expected profits equal the expected surplus from period-two innovation, i.e. $\bar{y}^*$ solves the following equation:

$$\int_{\bar{y}}^{\infty}[y+(1-\alpha)x]dG(y) + (1-\alpha)\int_{0}^{\bar{y}}ydG(y) = \int_{0}^{\infty}ydG(y) \qquad \text{III.1}$$

The first term on the LHS in III.1 is the expected profit accruing to period-two innovator in the event of a patentable period-two innovation being made, where $(1-\alpha)$ is the proportion of the bargaining surplus accruing to period-two innovator; the second term is the share of the expected bargaining surplus accruing to period-two innovator if the improvement is not patentable. The existence of a $\bar{y}^*$ that solves III.1 is easily established by writing the LHS of III.1 as

$$\int_0^{\infty} ydG(y)+(1-\alpha)x\left[1-G(\bar{y})\right]-\alpha\int_0^{\bar{y}} ydG(y) \qquad \text{III.2}$$

and noting that the second and third term should add up to 0, their sum ranging from $(1-\alpha)x$ to $-\alpha E(y)$ as $\bar{y}^*$ ranges from 0 to ∞.

What follows is perhaps the simplest formalization that captures some of the basic differences between research and development and that allows patents to convey useful information to firms.

III.3 The model

Firms take three key decisions:

(i) whether to pay a research fee, φ, that yields a probability p of producing a successful prototype of a new product/process;

(ii) to what extent, x, to develop the research prototype by expending resources in development inputs, v;

(iii) how much final output, Q, to produce.

As usual with multi-stage games that use a sub-game perfect equilibrium as solution concept, our three-stage game can be solved 'backwards', starting with the final stage.

Stage 3 : Production

Assuming that the research lottery has been entered and has resulted in success for k firms, the Cournot-Nash equilibrium level of output of a typical firm i, Q_i^*, can be defined in terms of investment in development inputs by firm i, v_i, and by all other k−1 firms, $v_{-i} \equiv (v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k)$, i.e., $Q_i^* \equiv Q_i^*(v_i, v_{-i})$ and the resulting gross profits (net profits plus development costs, v) can be written as $\gamma(v_i, v_{-i}; k)$. Well-behaved preferences for the final product are assumed to ensure that $\gamma(\cdot)$ is concave and increasing in v_i. Notice that the above formulation allows for more than one patent to be granted within the product/process class that defines "the" industry (i.e. $k \geqslant 1$).[2]

Stage 2 : Development

Whereas the choice between a single- or a multi-patent regime affects the outcome of the game through its impact on the production stage, the profit-maximizing level of development

---

[2] Consider the following parametrizations in obvious notation:

$$p(q_i+q_{-i}) = A - (q_i+q_{-i})$$
$$C_i(q_i, v_i) = [c_o - x_i(v_i)]q_i + v_i$$
$$A = c_o$$

Then it is simple to show that:

$$q_i^*(v_i, v_{-i}) = \frac{k}{k+1}\, x_i(v_i) - \frac{1}{k+1} \sum_{j\neq i}^{k} x_j(v_j)$$

$$\gamma(v_i, v_{-i}, k) = [\frac{k}{k+1}\, x_i(v_i) - \frac{1}{k+1} \sum_{j\neq i}^{k} x_j(v_j)]^2$$

expenditures depends also on whether patents are granted to research prototypes or to fully-developed products/processes. In the former case the identity of the successful firm(s) is known before resources are committed to development, whereas in the latter case a firm i succeeding at the research stage will set its $\mathbf{v}_i$ so as to maximize the **expected** (as opposed to actual) profits earned in the last stage.

In order to make the model both interesting and realistic, development costs are assumed to be a continuous function of the extent, $\mathbf{x}$, of the improvement of the product/process being developed, i.e., $\mathbf{x}_i = \mathbf{x}(\mathbf{v}_i)$.

Depending on whether $\mathbf{x}$ is interpreted as the size of the outward shift of the inverse demand curve as a result of development or as the downward shift of the marginal cost curve, the model can accommodate both "product" and "process" innovations.

Stage 1 : Research

By paying a research fee, φ, each firm buys a probability p of producing a successful research prototype.

A justification for this very simple characterization of research is that it points sharply to some of the key differences between research and development by taking them to extremes: while it is widely accepted that research is 'more uncertain' than development, here the latter is assumed to be altogether non-stochastic. Research output is also recognized to be more "discontinuous" than the improvements obtained through development: whereas the 'extent' of product (or process) improvements, $\mathbf{x}$, is a continuous function of development expenditures, $\mathbf{v}$, a strong threshold effect is assumed to operate in the case of research, in the sense that unless a fixed research fee, φ, is paid, no new ideas can be produced, with additional expenditures yielding no increases in the probability of producing a successful research prototype.

The following notation will prove useful later on: let $h(p, m, j)$ be the probability that $(j-1)$ out of $\mathbf{m}$ entrants succeed in producing a research prototype, or, alternatively, as the probability that, conditional on one firm out of $\mathbf{m}$ having succeeded in producing a prototype, j others also succeed, i.e.

$$h(p, m, j) \equiv \binom{m-1}{j} p^{j}(1-p)^{m-j-1}$$

It should be noted that, if only one patent is awarded for each class of new products/processes, success at the research stage does not guarantee the award of the patent; thus, if j inventors succeed at the research stage and each has a $1/j$ chance of being awarded the patent, the probability of a successful inventor being awarded a patent in an $\mathbf{m}$-firm industry will be

$$z(p, m) \equiv \sum_{j=0}^{m-1} h(p, m, j)\frac{1}{j+1}$$

III.4 A taxonomy of patent regimes

In the previous chapter we saw that patentability standards can be classified according to a **novelty** criterion, whereby a patent regime can be either **strict** – if it defines the product space so that only **one** patent is granted for any given class of products/ processes – or **permissive**, if **multiple** patents are allowed.

In this section, we shall consider an additional criterion to define patentability standards, namely the criterion of **industrial applicability**. In this we shall depart from traditional patent models which, because of their failure to distinguish between research and development, cannot address the question of **what** is being patented, which clearly lies at the heart of the problem of designing an optimal patent system. In fact, building on the model sketched in the previous section, we can classify patent regimes according as to whether patents are awarded to research prototypes (**ideas-based** regime) or to fully-developed products (**product-based** regime). Of course, these two patent regimes can be considered as the two extremes of a continuous spectrum that grades patent applications according to their degree of industrial applicability. The advantage of our binary classification scheme is that it is rich enough to address significant policy issues whilst being simple enough to be tractable and (potentially) implementable.

Combining the two criteria for patentability, namely novelty and industrial applicability, we can then examine each of the four possible patent regimes in the resulting 2×2 matrix:

		Novelty Criterion	
		Permissive (multi-patent)	Strict (single-patent)
Industrial Applicability Criterion	Ideas-based	PI	SI
	Product-based	PP	SP

Table III.1

Patent regimes classified according to criteria of novelty and industrial applicability

We are now in a position to define the profits accruing to a potential entrant under each of the four possible patent regimes; as one of the aim of the model is to determine the effects of alternative patent regimes on industry structure, we shall take the number of active firms as endogenous and determined by a zero-profit condition.

SI: Strict, Ideas-based patent regime

Under an SI regime, a single patent is awarded to the firm that produces a successful research prototype; if j firms are successful, then each has a probability of $1/j$ of being awarded the patent. The key feature of the SI regime is that the patentee can plan development expenditures under certainty. If **N** firms enter the industry, then each can expect a profit equal to:

$$E\{\pi^{SI}\} \equiv pz(p,N) \max_{v} \left[\gamma(v;1)-v\right] - \varphi \qquad \text{III.3}$$

As p is the probability of success, z the probability of being awarded the patent if successful at the research stage, and $\max[\gamma(v;1)-v]$ the profits from development, the first term in III.3 is expected profits from research, while the last, φ, is the cost of research.

SP: Strict, Product-based patent regime

Unlike the SI regime, under an SP regime firms that succeed at the research stage have to develop their prototypes before the identity of the (single) patentee is revealed, and thus the profit-maximizing amount of resources invested in development will depend on both the number of entrants, n, and on the probability of success, p; as a result the expected per-firm profits under an SP regime are given by

$$E\{\pi^{SP}\} \equiv p \max_{v} \left[z(p,n)\ \gamma(v;1)-v\right] - \varphi \qquad \text{III.4}$$

PI: Permissive, Ideas-based patent regime

Let $\pi_i(k)$ be the profits made by a typical firm i at a (symmetric) Nash equilibrium in a $(k+1)$-firm oligopoly where each firm chooses independently its own level of development expenditure, i.e.

$$\pi_i(k) \equiv \max_{v_i} [\gamma(v_i, v_{-i}, k+1) - v_i] \qquad \text{III.5}$$

Then

$$E\{\pi^{PI}\} \equiv p \sum_{k=0}^{N-1} h(p, N, k)\pi(k) - \varphi \qquad \text{III.6}$$

As p is the probability that a firm succeeds at the research stage and $h(p, N, k)$ is the probability that k out of the remaining $N-1$ firms are also successful, the first term in III.6 is expected profits from research, from which the research cost, φ, has to be subtracted. Notice that, as patents are awarded to research prototypes, firms commit resources to development expenditures knowing with certainty the structure of the final-output market.

PP: Permissive, Product-based patent regime

Under a PP regime, patents are granted to any firm that has both succeeded at the research stage and developed its prototype so as to meet the (more demanding) criterion of industrial applicability. As patents are now awarded to fully developed products/processes, those firms that have been successful at the research stage have to invest in development expenditures without knowing how many other firms have also succeeded, i.e. without knowing the structure of the final product market. As a result, the underlying game becomes one of *incomplete* information: the number of (active) players is unknown at the time when resources are committed to developing a successful prototype.

Taking Bayesian equilibrium as the appropriate solution concept for games with incomplete information, it is easy to see that the game that firms play under a PP regime is but a very simple example of a Bayesian game. As is well-known (see Fudenberg & Tirole (1986)), a Bayesian equilibrium is a straightforward extension of the Nash equilibrium concept in which each player recognises that the other players' strategies depend on their "types", or "characteristics". In our case at the beginning of the research stage each firm, upon payment of the research fee φ, knows its own type, i.e. whether it is "successful" or "unsuccessful", but, of course, does not know the type of all other firms. However each firm knows the probability distribution from which "types" are drawn and each firm knows that every firms knows it, etc. – the probability p is common knowledge.

Thus, it turns out that the underlying Bayesian game is exceedingly simple, not only because firms can belong only to either of two types, "successful" or "unsuccessful", but also because "unsuccessful" types simply do not participate to the development stage of the game. As a result, the typical maximand of a successful firm will be

$$E\{\pi^{PP}\} \equiv p \max_{v} \left[\sum_{k=0}^{n-1} h(p,n,k) \ \gamma(v;k) - v \right] - \varphi \qquad \text{III.7}$$

In III.7 $\mathbf{h(p, n, k)}$ is the probability that a successful firm be faced with $\mathbf{k}$ other successful inventors and $\boldsymbol{\gamma}\mathbf{(v, k)}$ is the **gross** profit accruing to the firm in a $\mathbf{(k+1)}$-firm oligopoly; thus $\sum \mathbf{h}(\cdot)\boldsymbol{\gamma}\mathbf{(v; k)}$ is the expected gross benefit from development.

III.5 Product- vs Ideas-based Single-patent Regimes

In terms of Table III.1 the previous chapter can be seen as an analysis of the welfare implications of the strictness (or, alternatively, the permissiveness) of the novelty criterion for patentability standards and its qualitative results apply irrespectively of whether the patent regime is ideas- or product-based.

In this chapter the emphasis is on the welfare implications of the industrial applicability criterion and in this section we shall consider whether it is preferable to grant patents to research prototypes or to fully developed products in a single-patent regime. In this respect we can prove the following:

Theorem III.1: Assuming that the number of firms can be treated as a continuous variable, a strict ideas-based regime is unambiguously welfare superior to a strict product-based regime (provided the latter sustains at least two firms).

Proof. Let $(1-p)\equiv q$, then using the identity[3]

$$\frac{(1-q^N)}{N} \equiv p \sum_{i=0}^{N-1} \binom{N-1}{i} p^i q^{N-i-1} \frac{1}{i+1}$$

III.3 and III.4 can be rewritten as:

$$E(\pi^{SI}) \equiv \left(\frac{1-q^N}{N}\right) \max_v \{\gamma(v)-v\} - \varphi \qquad \text{III.8}$$

$$E(\pi^{SP}) \equiv \left[\frac{1-q^n}{n}\right] \max_v \{\gamma(v) - [\frac{pn}{1-q^n}] v\} - \varphi \qquad \text{III.9}$$

Writing expected profits under the two regimes in the above form has the advantage of showing very clearly that under the SP regime it is as if firms faced a higher unit cost of development as compared with firms under an SI regime; in fact, for $\forall n \geq 2$, $\frac{pn}{1-q^n}$ (which can be interpreted as the unit cost of development under an SP regime) exceeds unity[4] (i.e. the unit cost of development under an SI regime). As a result, firms will invest less in development under an SP regime than under an SI regime. In fact if we let v^{SI} and v^{SP} be the solutions of the maximization of III.8 and III.9 respectively, i.e.,

$\gamma'(v^{SI}) = 1$ and $\gamma'(v^{SP}) = \frac{pn}{1-q^n}$ then as $\gamma(\cdot)$ is concave in v and $\frac{pn}{1-q^n} > 1$ (provided $n \geq 2$), the SI regime yields 'more improved' products/processes: $v^{SI} > v^{SP}$.

---

(3) $$\sum_{i=0}^{N-1} \binom{N-1}{i} p^{i+1} q^{N-(i+1)} \frac{1}{i+1} \equiv \sum_{i=1}^{N} \binom{N-1}{i-1} p^i q^{N-i} \frac{1}{i} \equiv$$
$$\equiv \sum_{i=1}^{N} \binom{N}{i} \frac{i}{N} p^i q^{N-i} \frac{1}{i} \equiv \frac{1-q^N}{N}$$

(4) $$\frac{1-q^n}{p} \equiv \sum_{i=1}^{n} \binom{n}{i} p^{i-1} q^{n-i} \equiv \sum_{i=0}^{n-1} \binom{n}{i+1} p^i q^{n-i-1} \equiv$$
$$\equiv \sum_{i=0}^{n-1} \binom{n-1}{i} \frac{n}{i+1} p^i q^{n-i-1} \equiv n \sum_{i=0}^{n-1} \binom{n-1}{i} \frac{n}{i+1} p^i q^{n-i-1} < n.$$

In line with much of the literature on the subject, the criterion deployed here to assess which patent regime is socially superior is net expected social welfare, E{W}, defined as the sum of expected consumer surplus CS and industry profits Π. The general formula for E{W} is thus:

E{W} ≡ [probability of at least one discovery] × [gross social welfare under certainty] - [expected industry development costs] - [industry research costs]

Applying the above definition to the SI and SP regimes and defining $G(v)$ as gross industry profits, we obtain

$$E\{W^{SI}\} \equiv (1-q^{N}) \; \{CS(v^{SI})+G(v^{SI})-v^{SI}\}-N\varphi \qquad \text{III. 10}$$

$$E\{W^{SP}\} \equiv (1-q^{n}) \; \{CS(v^{SP})+G(v^{SP})\}-npv^{SP}-n\varphi \qquad \text{III. 11}$$

If the integer constraint is ignored and research costs, φ, are such that $E\{\pi^{SI}\}=E\{\pi^{SP}\}=0$ with both $\mathbf{N}$ and $\mathbf{n}$ being integers then it is easy to show that

$$E\{W^{SI}\}\equiv(1-q^{N}) \; CS(v^{SI})>(1-q^{n}) \; CS(v^{SP})\equiv E\{W^{SP}\} \qquad \text{III. 12}$$

The above inequality follows by noting that as consumer surplus is increasing in $\mathbf{v}$, $CS(v^{SI})>CS(v^{SP})$ and as $\gamma(\cdot)$ is increasing in $\mathbf{v}$, the two regimes cannot sustain the same number of firms in a zero-profit equilibrium, for if $\mathbf{N} = \mathbf{n}$ then $E\{\pi^{SI}\} > E\{\pi^{SP}\}$. Then, for both $E\{\pi^{SI}\}$ and $E\{\pi^{SP}\}$ to be equal to zero, $\mathbf{N}$ must exceed $\mathbf{n}$, for $(1-q^{k})/k$ is decreasing in

$k^{(5)}$. ■

Theorem III.1 says that, if the integer constraint is ignored, then there is no benefit to be had by generating uncertainty with the introduction of a product-based patent regime: as a result of the additional cost engendered by uncertainty (i) firms will invest less in development and thus the lucky patentee-monopolist will restrict output to a larger extent than under an SI regime; and (ii) fewer firms will enter the research stage, thereby reducing the probability of at least one successful research prototype being produced at all.

However, in the almost certain event of either **N** or **n** (or both) not being integers, the question must be asked whether the integer constraint can reverse the above inequality. Obviously, if **n** were an integer but not **N**, the case for a SI regime would be strengthened (as the super-normal profits enjoyed by the [N] firms would have to be added to consumers surplus, where [N] is the nearest integer less than **N**). To ascertain whether a non-integer **n** can be of qualitative consequence, one has to

---

(5) $\frac{1-q^k}{k}$ is decreasing in k iff

$$\frac{1-q^k}{k} > \frac{1-q^{k+1}}{k+1} \quad \text{i.e.}$$

$$1 > q^k(1-pk) \qquad \text{III. 13}$$

As the r.h.s. of III.13 reaches a maximum at p^*:

$$(1-p^*)^{k+1} = (1+p^*k)(1-p^*)^k \qquad \text{III. 14}$$

substituting III.14 into III.13, it can be seen that inequality III.13 holds, for

$$1>(1-p^*)^{k+1}. \quad ■$$

resort to some parametrization. For ease of comparison with the rest of the literature, I shall use the popular parametrization of a linear inverse demand curve, constant marginal costs and an iso-elastic development function, i.e.

$$P = P_o - \sum_{i=1}^{m} Q_i \; ; \qquad m = n, \; N \qquad \text{III.15}$$

$$C = (c_o - x_i)Q_i + v_i \qquad \text{III.16}$$

$$x_i = \theta v_i^{\alpha} \qquad \text{III.17}$$

The above assumes that improvements brought about by development activity take the form of a downward shift of pre-innovation marginal cost; equivalently it can be assumed that development shifts the pre-innovation inverse demand curve by x (i.e., $P = P_o + x - \sum_{i=1}^{m} Q_i$).

For simplicity, it is also assumed that without the innovation production is marginally unprofitable, i.e., $c_o = P_o$.

Straightforward substitution yields

$$v^{SI} = \left(\frac{\alpha\theta^2}{2}\right)^{1/(1-2\alpha)} \qquad \text{III.18}$$

$$\gamma(v^{SI}) = \frac{1}{2\alpha} v^{SI} \qquad \text{III.19}$$

$$v^{SP} = \left(\frac{1-q^n}{np}\right)^{1/(1-2\alpha)} v^{SI} \qquad \text{III.20}$$

$$\gamma(v^{SP}) = \frac{1}{2\alpha} \frac{np}{1-q^n} v^{SP} \qquad \text{III.21}$$

Taking into account that $CS(v) = \frac{1}{2}G(v)$, substituting III.18-21 into III.10-11 and assuming that research costs are such that $E\{\pi^{SI}\} = 0$ with N being an integer, we finally obtain

$$E\{W^{SI}\} = (1-q^{N})v^{SI}/4\alpha \qquad \text{III.22}$$

$$E\{W^{SP}\}=(1-q^{[n]})3\gamma(v^{SP})/2-[n]pv^{SP}-[n](1-2\alpha)(1-q^{N})v^{SI}/2\alpha N \qquad \text{III.23}$$

where $[n]$ is the nearest integer less than n.

Thus $E\{W^{SI}\} \gtrless E\{W^{SP}\}$ according as to whether

$$(1-q^{N})\frac{[N+2[n](1-2\alpha)]}{N} \gtrless [n]p(3-4\alpha)\left(\frac{1-q^{[n]}}{[n]p}\right)^{1/(1-2\alpha)} \qquad \text{III.24}$$

Inequality III.24 does not lend itself to simple generalizations about the qualitative effects of changes in the parameters on expected social welfare. However, some general points can be made:

(i) There exist combinations of p and α such that the introduction of the integer constraint does reverse the general presumption of the welfare superiority of the SI regime (see Theorem III.1). The results of simulations with a range of values of the crucial parameters are reported in Table III.2.

(ii) As shown in Table III.2, as inventions become more difficult, i.e. as p falls, the SI regime becomes correspondingly more attractive and in the limit (as $p\to 0$) is unambiguously superior to the SP regime[6].

---

[6] Writing III.24 as

$$\left(\frac{1-q^{N}}{p}\right)[N+2[n](1-2\alpha)]/N \gtrless [n](3-4\alpha)\left(\frac{1-q^{[n]}}{[n]p}\right)^{1/(1-2\alpha)}$$

and applying de l' Hospital's rule to both sides we obtain

$$\lim_{p\to 0}\left(\frac{1-q^{N}}{p}\right) = N\ ;\ \lim_{p\to 0}\left(\frac{1-q^{n}}{np}\right)^{1/(1-2\alpha)} = 1$$

and thus $\lim_{p\to 0} E\{W^{SI}\} > \lim_{p\to 0} E\{W^{SP}\}$, as $N>n$.

The economic rationale for this result is easy to appreciate: with inventions being "difficult", the super-normal profits generated by a non-integer **n** are insufficient to offset the benefit generated by the larger number of inventors sustained by the SI regime.

(iii) The converse does not hold, i.e. as **p** approaches unity, the SP regime does <u>not</u> become necessarily preferable to the SI regime. One of the reasons behind this asymmetry is rather intriguing; it can be shown that whilst $E\{W^{SI}\}$ is monotonically increasing in **p**, $E\{W^{SP}\}$ is not[7]. This may bring about a re-switching with the SI regime being welfare superior for values of **p** around **0** and **1** and the SP regime being preferable for intermediate values of **p**; Fig. III.1 illustrates this possibility.

---

[7] Ignoring the cost of research, $n\varphi$ (which is independent of p) and using, the parametrization III.15-17, $E\{W^{SP}\}$ can be written in the form

$$W = h\ p^{-2\alpha/(1-2\alpha)}\ (1-q^n)^{1/(1-2\alpha)} \text{ and thus}$$

$$\frac{dW}{dp} = \frac{W}{1-2\alpha}\left\{\frac{n(1-p)^{n-1}}{1-(1-p)^n} - \frac{2\alpha}{p}\right\}$$

Hence **W** has a turning point at $0<p^*<1$ and is negatively sloped at $p=1$; in fact at $p=1$ $dW/dp=-2\alpha$ and at $p=0$ $dW/dp=+\infty$ as can be shown by writing

$$\frac{dW}{dp} = \frac{W}{1-2\alpha}\left\{\frac{pn(1-p)^{n-1}-[1-(1-p)^n]2\alpha}{p[1-(1-p)^n]}\right\}$$

and applying de l' Hospital's rule.

Table III.2

Simulation results of welfare comparisons between SI and SP patent regimes

[Key to the table; bracketed (unbracketed) entries specify the range of values of p such that the *SP* (alt., SI) regime yields a higher expected net welfare level]

$\alpha=0,05$

N \ n	1	2	3	4	5	6	7
2	0,00-0,52 *[0,53-1,00]*						
3		0,00-0,36 *[0,37-1,00]*					
4			0,00-0,31 *[0,32-1,00]*				
5				0,00-0,32 *[0,33-0,85]* 0,86-1,00			
6					0,00-1,00		
7					*[0,67-1,00]*	0,00-0,66	
8						*[0,45-0,80]* 0,81-1,00	0,00-0,44

$\alpha=0,1$

N \ n	1	2	3	4	5	6
2	0,00-0,55 *[0,56-1,00]*					
3		0,00-0,48 *[0,49-1,00]*				
4			0,00-1,00			
5			*[0,51-1,00]*	0,00-0,50		
6				0,00-1,00		
7				*[0,77-0,84]* 0,85-1,00	0,00-0,76	
8					0,50-1,00	0,00-0,49

Table III.2 (continued)

$\alpha=0,15$

N \ n	1	2	3	4	5
2	0,00-0,58 *[0,59-1,00]*				
3		0,00-1,00			
4		0,41-0,50 *[0,51-1,00]*	0,00-0,40		
5			0,00-1,00		
6			*[0,60-0,63]* 0,61-1,00	0,00-0,59	
7			0,93-1,00	0,24-0,92	0,00-0,23
8				0,58-1,00	0,00-0,57

$\alpha=0,175$

N \ n	1	2	3	4	5
2	0,00-0,60 *[0,61-1,00]*				
3		0,00-1,00			
4		0,00-0,62 *[0,63-0,90]* 0,91-1,00			
5		*[0,86-1,00]*	0,00-0,85		
6			0,36-1,00	0,00-0,35	
7			0,69-1,00	0,00-0,68	
8			0,91-1,00	0,38-0,9	0,00-0,37

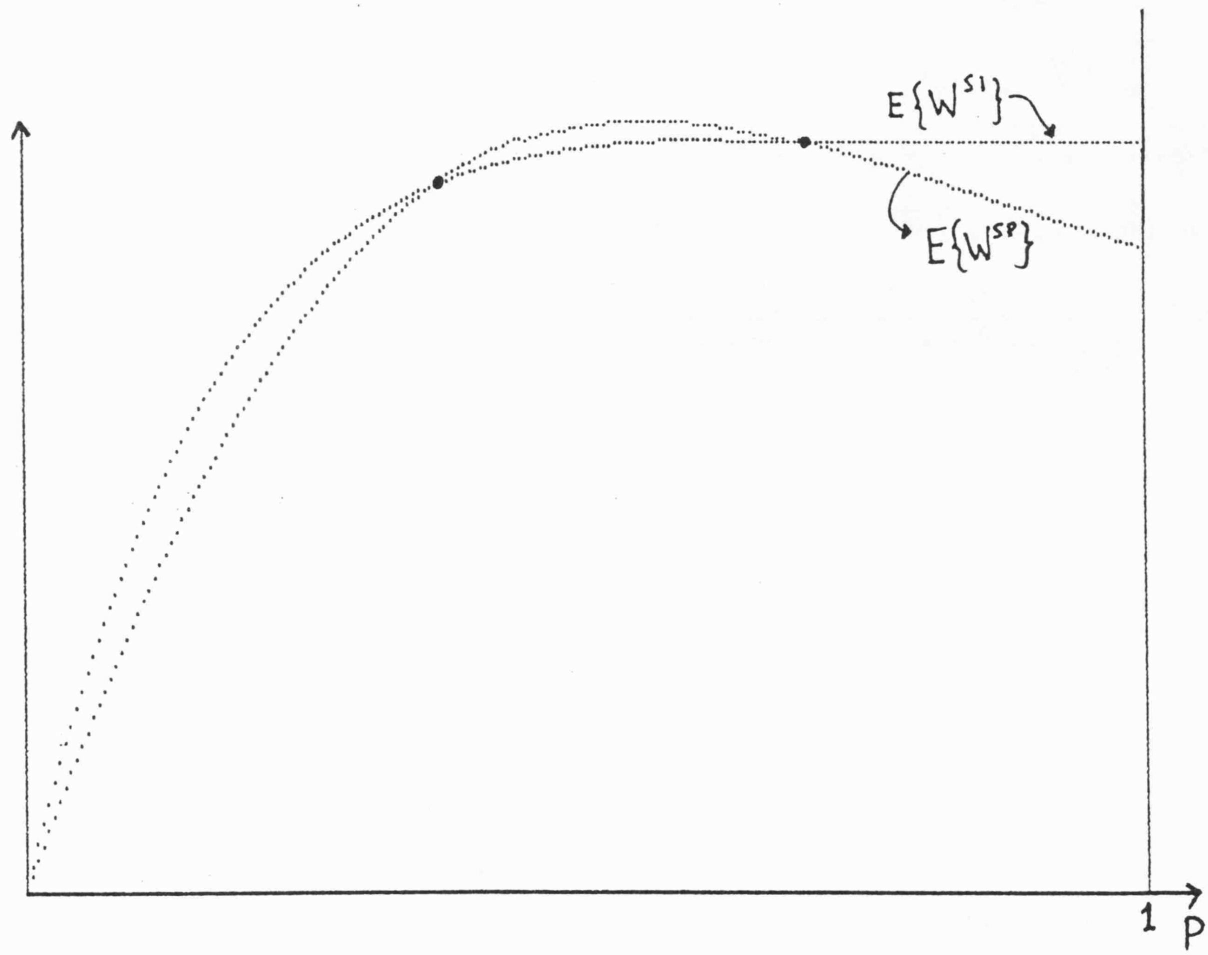


FIG. III.1

Re-Switching in an Optimal Strict Patent Regime

III.6 The Welfare Characteristics of Permissive Patent Regimes

Before comparing an ideas-based to a product-based mode under a permissive patent regime, I shall examine in some detail the welfare characteristics of permissive regimes, taking the PI case as an example. This choice is motivated by the fact that the PI regime can be seen as the opposite of the traditional winner-takes-all model in which no distinction is drawn between R and D (i.e., in our terminology, a SP regime). In fact, in the SP case the value of the award – the patentee's monopoly profits – does not depend on the number of entrants, but the probability of success does; on the contrary, under a PI regime the probability of discovery, p, is independent of market structure but the value of the prize is not.

In what follows I shall make extensive use of the parametrization III.15-17, not only because it serves well my purpose of illustrating the logic of the argument in a simple way but also because in many instances no general results are available and thus it is useful to have a feel for the contrasting forces at work.

Using III.15-17, it is straightforward to confirm that at a Nash equilibrium expected per-firm profits $E\{\pi^{PI}\}$, development expenditure v^{PI}, and per-firm output Q^{PI} can be written as

$$E(\pi^{PI}) = \sum_{i=1}^{N} \binom{N}{i} p^i q^{N-i} v_i^{PI} \frac{(1-2\alpha i)}{2\alpha N} - \varphi \qquad \text{III.25}$$

$$v_i^{PI} = \left(\frac{2 i \alpha \theta^2}{(i+1)^2}\right)^{1/(1-2\alpha)} \qquad \text{III.26}$$

$$Q_i^{PI} = \frac{\theta}{(i+1)} \left[v_i^{PI}\right]^{\alpha} \qquad \text{III.27}$$

Proposition III.1: Let $\varphi(p,N)$ be the cost of research such that in a free-entry equilibrium all N entrants earn zero profits (with N integer); then for any $N \geqslant 2$ $\varphi(p,N)$ reaches a maximum at $\hat{p} \in (0,1)$.[8]

---

[8] As $\quad \varphi(p,N) \equiv \sum_{i=1}^{N} \binom{N}{i} p^i q^{N-i} \max_{v_i} \{\gamma(v_i, v_{-i}) - v_i\}$

it is easy to establish that there exists a value of $p \in (0,1)$, $\hat{p}$, such that:

$$\left.\frac{d\varphi(p,N)}{dp}\right|_{p=\hat{p}} = 0 \qquad \text{III.27}$$

Differentiating $\varphi(p,N)$ w.r.t. p we obtain

$$\frac{d\varphi(p)}{dp} = q^{N-1}\Pi_1 + \sum_{i=1}^{N-1} \binom{N-1}{i} p^i q^{N-1-i} [\Pi_{i+1} - \Pi_i]$$

where $\Pi_i \equiv i\,[\max_{v_i} \{\gamma(v_i, v_{-i}) - v_i\}] \equiv i\pi_i$.

As $d\varphi(p)/dp$ is continuous and its values at p=0 and p=1 are respectively $\Pi_1 > 0$ and $(\Pi_N - \Pi_{N-1}) < 0$, there exists a $\hat{p} \in (0,1)$ such that III.27 is satisfied.

Let $\sigma_1 \equiv \Pi_1 - \Pi_2 > 0$, then straightforward manipulation yields that at $p=\hat{p}$

$$\text{sign}\frac{d^2\varphi(p,N)}{dp^2} = \text{sign}\left\{-q^{N-1}(N-1)\sigma_1 - \sum_{i=1}^{N-2} \binom{N-1}{i} p^i q^{N-1-i} (N-1-i)\pi_{i+1}\right\} < 0$$

thus proving that $\varphi(p,N)$ is dome-shaped with a maximum at $\hat{p} \in (0,1)$. ∎

The fact that for any $N\geq 2$ $\varphi(p,N)$ is dome-shaped may seem puzzling: given a certain level of research expenditure one would expect that as discoveries become more difficult (i.e. as p falls) the number of firms sustained by a free-entry equilibrium should fall. Proposition III.1 asserts that there is always a level of research fee φ^{*}, such that $\varphi(\tilde{p},N)>\varphi^{*}>E\{\pi^{FI}(p=1;N)\}$. This means that for $\varphi=\varphi^{*}$ there is always a reswitching in the equilibrium number of firms as p varies. Fig. III.2 illustrates this case for $N=2$.

The economic rationale for a re-switching is simple. Suppose that discovery is easy (i.e. $p^{H}<p\leq 1$) and research inputs are so expensive that a free-entry equilibrium sustains only one firm. As discoveries become slightly more difficult but still "easy" (i.e., $p^{M}\leq p\leq p^{H}$) two firms may find entry profitable; their profits, of course, will be a weighted average of monopoly and duopoly profits. As the weight attached to monopoly profits $(p(1-p))$ is small, duopoly profits may be large enough to make entry actually profitable for two firms; this is because for $p^{H}<p\leq 1$ a free-entry monopolist earns substantial super-normal profits. As p falls, expected profits also fall, but it is expected monopoly profits that become relatively more important. Eventually a point will be reached (p^{L} in Fig. III.2) where a free-entry equilibrium will again sustain only one firm.

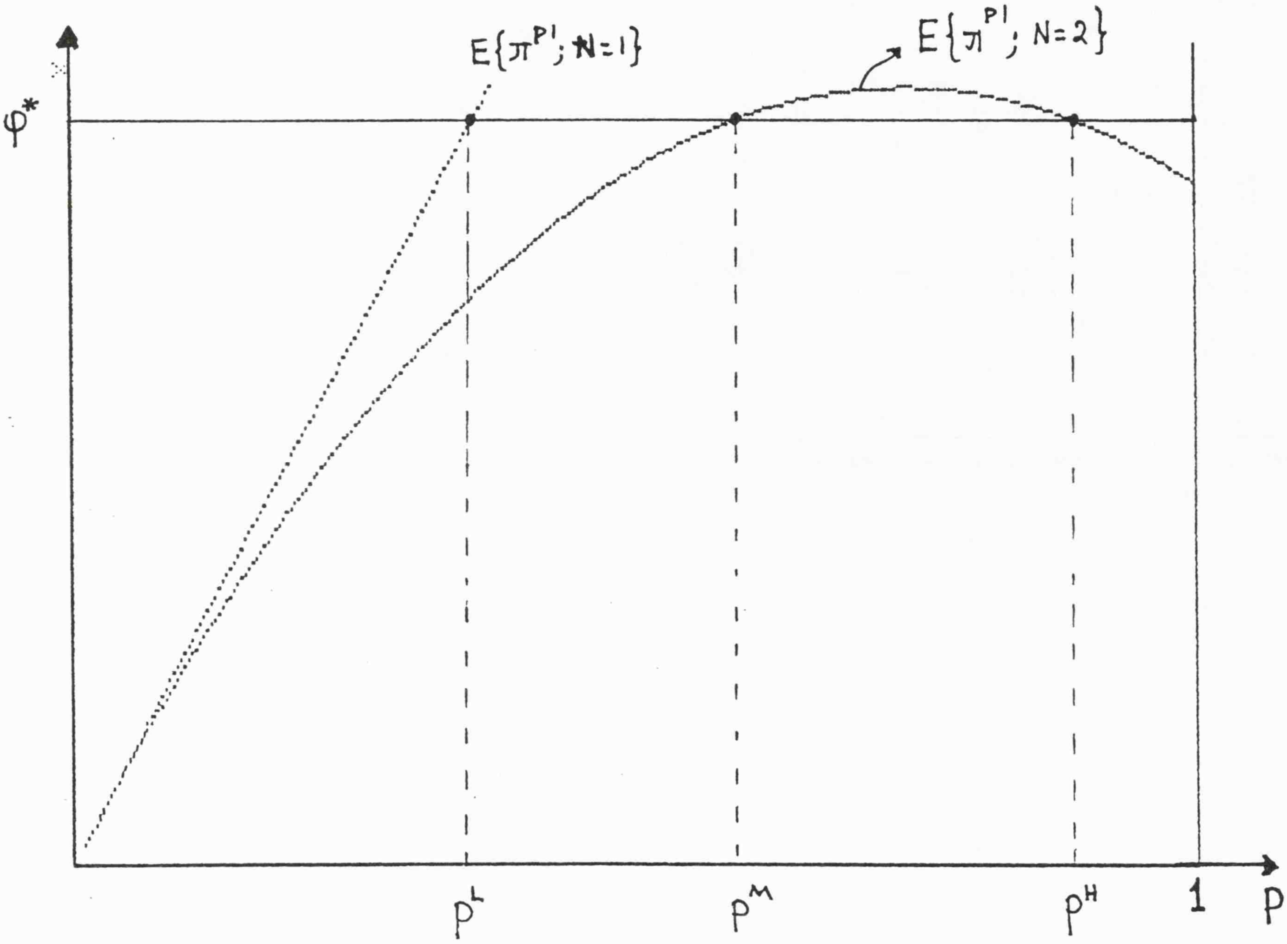


FIG. III.2

A re-switching in the free-entry number of firms under a permissive ideas-based patent regime

N=1 for $p^H < p \leq 1$

N=2 for $p^M \leq p \leq p^H$

N=1 for $p^L \leq p \leq p^M$

N=0 for $p < p^L$

Proposition III.2: Let $\hat{p} \equiv \operatorname{argmax} \varphi(\hat{p}, N)$ and $\hat{p}^* \equiv \operatorname{argmax} \varphi(\hat{p}^*, N+1)$;

then $\hat{p} > \hat{p}^*$.[9]

Proposition III.2 states that the level of probability p at which expected gross profits reach a maximum, i.e., $\hat{p}$, falls as the number of entrants **N** grows.

The significance of Propositions III.1 and III.2 can be be appreciated by giving a slightly different interpretation to the above model. Suppose that **N** firms, which behave non-cooperatively in their final product market, finance a jointly-owned research facility, which can produce two types of reserach prototypes: perfectly reliable prototypes, that can be developed with certainty into marketable products, and imperfectly reliable prototypes, that can be turned into saleable products with probability $p<1$. then Proposition III.1 states that all **N** firms will agree on producing the less reliable prototype; and according to Proposition III.2 the larger is the number of firms funding the jointly-owned research laboratory, the less reliable

---

[9] To prove Proposition III.2 it suffices to show that

$$\left.\frac{d\varphi(p, N+1)}{dp}\right|_{p=\hat{p}} < 0 \quad \text{where } \hat{p} \text{ is defined by}$$

$$(1-\hat{p})^{N-1}\pi_1 + \sum_{i=1}^{N-1}\binom{N-1}{i}\hat{p}^i(1-\hat{p})^{N-1-i}(\Pi_{i+1}-\Pi_i)=0 \qquad \text{III.29}$$

using the same notation of III.28.
Multiplying III.29 by $(1-\hat{p})$ and substituting it into $\frac{d\varphi(p, N+1)}{dp}$ we obtain

$$-\sum_{i=1}^{N-1}\binom{N-1}{i}\hat{p}^i(1-\hat{p})^{N-i}(\Pi_{i+1}-\Pi_i)+\sum_{i=1}^{N-1}\binom{N}{i}\hat{p}^i(1-\hat{p})^{N-i}(\Pi_{i+1}-\Pi_i)+\hat{p}^N(\Pi_{N+1}-\Pi_N)$$

$$\equiv \sum_{i=1}^{N-1}\left[\binom{N}{i}-\binom{N-1}{i}\right]\hat{p}^i(1-\hat{p})^{N-i}(\Pi_{i+1}-\Pi_i)+\hat{p}^N(\Pi_{N+1}-\Pi_N)$$

$$\equiv \sum_{i=1}^{N-1}\binom{N}{i}\frac{i}{N-i}\hat{p}^i(1-\hat{p})^{N-i}(\Pi_{i+1}-\Pi_i)+\hat{p}^N(\Pi_{N+1}-\Pi_N) < 0. \quad \blacksquare$$

will be the profit-maximizing prototype. Consider the case of N=2 and let π_1 and π_2 be profit per-firm under monopoly and duopoly, respectively; then Proposition III.1 can be interpreted as saying that:

$$p(1-p)\pi_1 + p^2\pi_2 > \pi_2 \qquad \text{III.30}$$

i.e., letting $\pi_1 \equiv (1+k)\pi_2$, with $k>1$ to allow for the dissipation of profits engendered by competition

$$pk > 1 \qquad \text{III.31}$$

Maximizing the l.h.s. of III.30 w.r.t. p yields

$$\hat{p} = \frac{1+k}{2k} < 1 \qquad \text{III.32}$$

thus the two firms will instruct the research laboratory to produce a research prototype of **imperfect** reliability. More perversely still, the same result obtains even if producing an imperfectly reliable prototype is (slightly) **more** costly than producing one of perfect reliability. Proposition III.2 may be interpreted as a warning against a possible side-effect of lowering the cost of research (e.g. by subsidizing international research consortia), in so far as the attending increase in membership in the joint research facility yields research prototypes of decreasing reliability(10).

---

(10) The above example can also be used to throw a different light on the frequently-heard complaint about the unsatisfactory level of industrially-applicable research in the UK. If one considers the education system as the nation's 'research laboratory' and British industry as a (admittedly indirect) determinant of the quality of its research output (e.g. by giving low status and salaries to applied

As a preliminary step towards examining the welfare characteristics of a PI patent regime with free entry under uncertainty, it may be useful to analyse the simpler case of a permissive regime with exogenously given market structure under certainty (notice that under certainty the distinction between ideas- and product-based regimes disappears).

In particular, it is interesting to examine the effects of changes in the number of active firms. Increased entry (in a comparative-statics sense) produces three effects:

(i) **higher industry-wide research costs:** in the absence of uncertainty, the research fee incurred by an additional entrant yields no social benefits;

(ii) **lower rate of technical change:** i.e., $\frac{dv}{dN} < 0$; cost reductions are lower, the larger the number of active firms.

(iii) **larger output:** $\frac{dNQ}{dN} > 0$; entry (provided it is feasible) will always raise industry output[11].

---

scientists), then the above model suggests that resulting low percentage of research output turned into marketable products/ processes can in fact be the direct result of profit maximization. Notice that, albeit under somewhat stricter conditions, the same perverse results obtain even if an outsider (a Japanese company?) is committed to perfectly reliable research prototypes.

[11] Using III.22-25 it is easy to check that in order for entry to be profitable for the (N+1)th firm $\frac{1}{2(N+1)} > \alpha$ industry output rises with entry iff $\frac{1}{N+1} > \alpha$.

The fact that opposing forces are at work raises the obvious question as to whether entry may reduce welfare.

Using the parametrization III.15-17 it is easy to confirm that for any given N, aggregate social surplus under a permissive patent regime with no uncertainty is given by

$$W(N) \equiv \frac{N(1-4\alpha)+2}{4\alpha} v(N) - N\varphi \qquad \text{III. 33}$$

where (see III.26)

$$v(N) = \left(\frac{2N\alpha\theta^2}{(N+1)^2}\right)^{1/(1-2\alpha)} \qquad \text{III. 34}$$

Define $\varphi^*(N,\alpha)$ and $\varphi^\circ(N,\alpha)$ respectively as the research fee such that the net social benefit of the Nth firm is zero and the maximum research fee that sustains an N-firm oligopoly, i.e.

$$\varphi^*(N,\alpha) \equiv \frac{N(1-4\alpha)+2}{4\alpha} v(N) - \frac{(N-1)(1-4\alpha)+2}{4\alpha} v(N-1) \qquad \text{III. 35}$$

$$\varphi^\circ(N,\alpha) \equiv \frac{1-2\alpha N}{2\alpha N} v(N) \qquad \text{III. 36}$$

It is straightforward to confirm that the relative positions of $\varphi^*(N,\alpha)$ and $\varphi^\circ(N,\alpha)$ in the (φ,α) plane are as depicted in Fig. III.3.

From Fig. III.3 it also transpires that there is a range of values of (φ,α) such that entry (in a comparative-statics sense) may both **feasible** and **socially detrimental**. For instance, for all pairs (φ,α) within the shaded area of Fig. III.3, the benefit of additional output brought about by a third entrant is more than offset by the extra research cost.

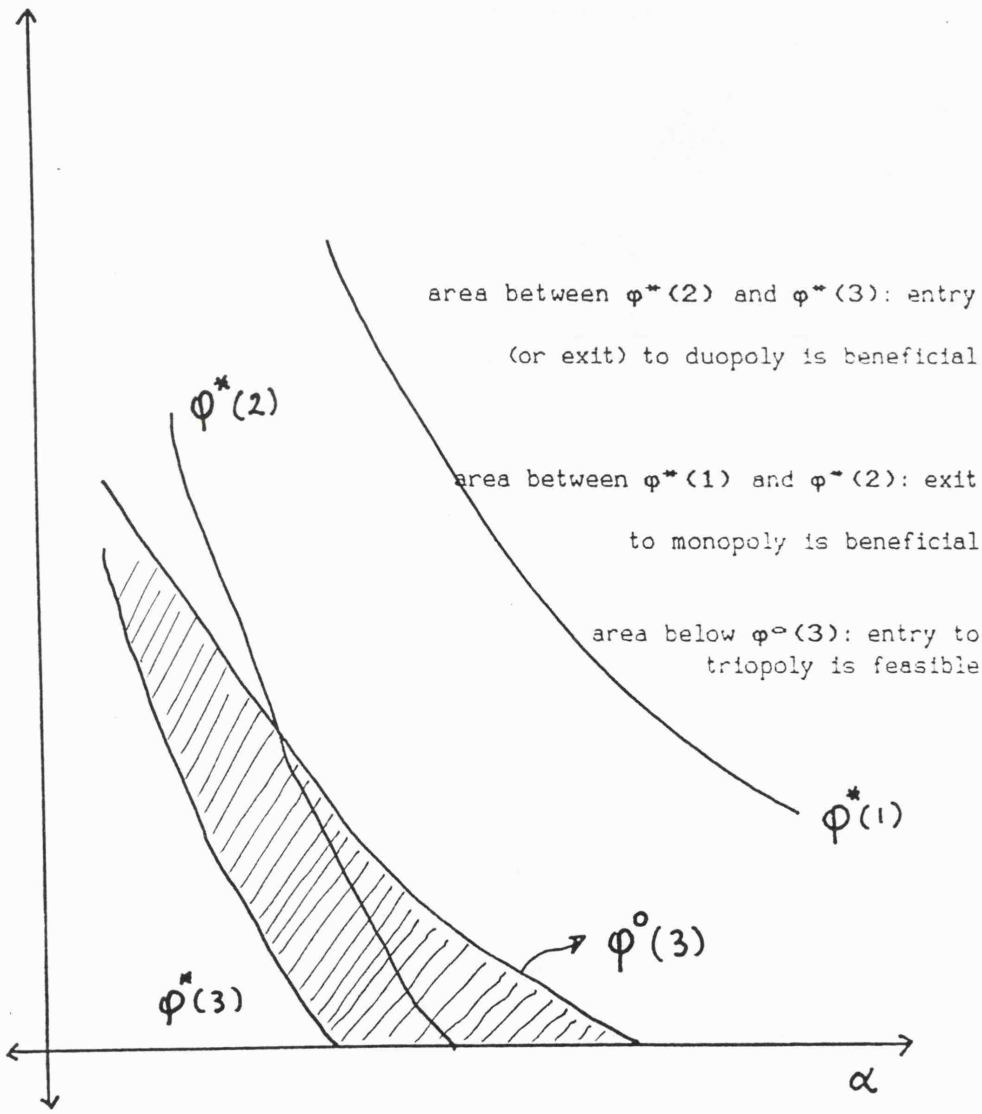


FIG. III.3

Entry both feasible and socially detrimental

Finally, we can now introduce an endogenously determined number of firms and uncertainty. Tedious substitution reveals that for the parametrization III.15-17, net expected social welfare can be written as

$$E\{W^{PI}(N)\} = \sum_{i=1}^{N} \binom{N}{i} p^i q^{N-i} v_i^{PI} \frac{2+i(1-4\alpha)}{4\alpha} - N\varphi \qquad \text{III. 37}$$

Even by resorting to the parametrization III. 15-17 $E\{W^{PI}\}$ does not simplify to a well-behaved function; however, some interesting results can be obtained by comparing the welfare characteristics of a N-firm free-entry oligopoly as compared to a (N-1)-firm oligopoly. For this purpose define

$$\Delta(p, N) \equiv E\{W^{PI}(N)\} - E\{W^{PI}(N-1)\} \qquad \text{III. 38}$$

Obviously for any given N and α $\Delta(p, N)=0$ defines all the pairs (φ, p) such that the marginal expected gross social benefit derived from the Nth firm is equal to φ, the cost of research.

We are interested in the relationship between $E\{\pi^{PI}(N)\}=0$ and $\Delta(p, N)=0$; the former determines the number of firms active at a free-entry equilibrium and the latter helps establish whether such equilibrium sustains too few or too many firms. It should be clear that, unlike the traditional winner-takes-all patent race (in which social welfare is monotonically decreasing in the number of entrants), under a PI regime a free-entry equilibrium may well sustain **too few** entrants. This is because increased entry has three conflicting effects on research and development: if discoveries are "difficult" and research costs are not "too" high, a large number of research units may be beneficial in so far as it generates a high probability of a discovery being made at all; however, a large number of successful inventors will produce two further effects: on the one hand it will yield a high level of final output, but on the other will generate a low level of cost reduction (or product improvement). By balancing these three effects one can define the second-best number of firms, N^s, which need not coincide with the free-entry equilibrium number of entrants, N^o.

Figure III.4 depicts $\Delta(p, N)=0$ and $E\{\pi^{PI}(N)\}=0$ for $\alpha=0.1$ (the average value of the development cost elasticity according to the empirical literature on R&D); it can be seen that the (φ, p) plane can be divided into three zones:

ZONE A: $N^s>N^o$. When discoveries are "difficult" relatively to the cost of research, a free-entry equilibrium does not generate enough research units.

ZONE B: $N^s=N^o$. Social welfare cannot be improved by imposing restrictions to entry.

ZONE C: $N^s<N^o$. Although the rugged contours of Zone C do not allow us to draw clear-cut conclusions, it can be seen that when discoveries are "easy" (i.e. p is "high") and the cost of research "low", a free-entry equilibrium sustains ***too many firms*** (and thus any policy change that makes entry less profitable may improve welfare).

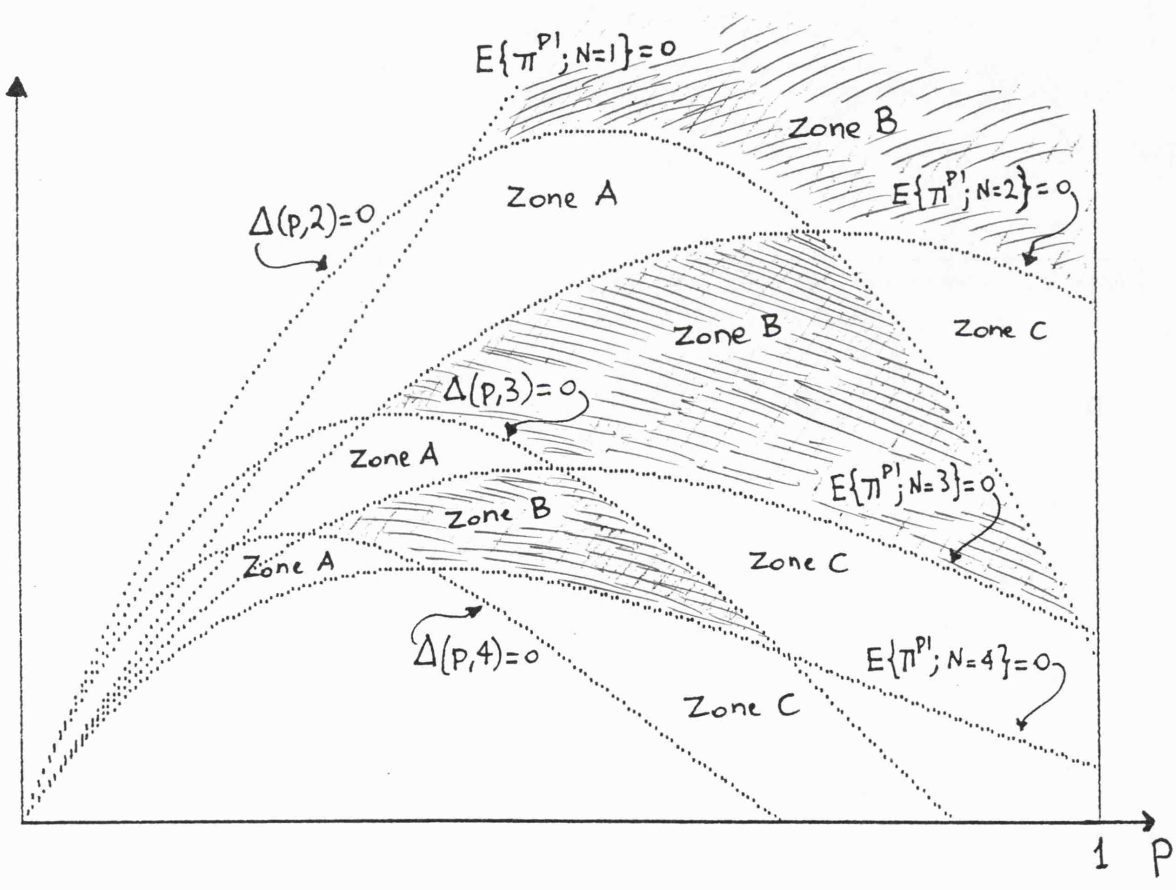


FIG. III.4

Second-best and Free-entry Industry Structure under a Permissive Ideas-Based Patent Regime

It is easy to show that

Proposition III.3: Under certainty, given the parametrization III.15-17, a decrease in the free-entry number of firms is always welfare-improving.

Let N^o+1 be the free-entry number of firms, i.e., recalling III.35,

$$\frac{1-2\alpha(N^o+1)}{2\alpha(N^o+1)} = \varphi \qquad \text{III. 39}$$

Using III.34 and III.39, $W^{PI}(N)>W^{PI}(N^o+1)$ can be written as

$$\frac{N^{o2}+3N^o+2-4\alpha(N^o+1)}{N^{o2}+4N^o+1-4\alpha(N^o+1)} > \frac{v^{PI}(N^o+1)}{v^{PI}(N^o)} \equiv \left\{\frac{(N^o+1)^3}{N^o(N^o+2)^2}\right\}^{1/(1-2\alpha)} \qquad \text{III. 40}$$

Taking into account that the r.h.s. of III.40 is largest at $\alpha=0$ and that setting $\alpha=½$ lowers the l.h.s., the above inequality will certainly hold if

$$\frac{N^o}{N^{o2}+2N^o-1} \geq \frac{(N^o+1)^2}{N^o(N^o+2)^2} \quad \text{which holds for } \forall N^o. \quad \blacksquare$$

III.7 Patents as Early Warning Devices

The main result of the previous section is that, especially when discoveries are "easy" as compared to research costs, a permissive ideas-based patent regime may generate **excessive research** and **too little development**.

The rationale for this result (and a possible suggestion for improving social welfare by a change of patent regime) can be understood by viewing patents as **information signals**. Unlike the traditional model of patent races which, by defining patents as the reward of the **combined** R&D effort, deprives the patent system of any informational role, in our model patents can alter the allocation of resources **between** research and development by taking on the role of "early warning devices" in the sense that by a suitable choice of patentability criteria (i.e. ideas- or product-based) patents may be used either to convey or to withhold valuable information before resources are committed to developing research prototypes into finished products/ processes.[12]

We are now in a position to compare the two possible patent regimes under a permissive mode; it should be clear that in assessing the relative merits of an ideas-based vs a product-based regime a key role is played by uncertainty. In fact, unlike the case of strict regimes where the distinction between ideas- and product-based patentability criteria does not vanish

---

[12] A distinction should be drawn between our model and Kitch's views on the nature of the patent system. It will be recalled that Kitch (1977) argues that patents are analytically similar to "mineral claims" (i.e. patents are viewed are "prospects") and are socially valuable not so much as a reward for inventing but rather as an efficient means of organizing post-invention activities. The argument developed here is totally immune from any reference to transaction costs and treats patentability standards (i.e. whether patents are "rewards" or "prospects") as policy instruments to be used so as to affect the allocation of resources between research and development in conditions of uncertainty.

even under certainty, in conditions of certainty the two permissive regimes are identical. If successful research prototypes can be produced under certainty, it makes no difference whether patents are awarded to prototypes or to fully developed products/processes, for in either case the structure of the final product industry is known before resources are committed to development.

Again unlike the case of single-patent (i.e. strict) regimes, in a multi-patent system the question of whether welfare would be higher under an ideas- or product-based regime is bound to be an open one, for the effects of entry are quite different: whereas under a strict system an ideas-based regime has the unambiguous advantage of inducing more entry (and thus a higher probability of at least one successful prototype being produced at all) with no adverse effects on the extent to which the product/process is developed, under a multi-patent system, increased entry is not necessarily welfare-improving. In fact, the larger number of entrants sustained by an ideas-based regime implies that patentees would develop their prototypes less than if there had been fewer entrants. The question may then be asked of whether a change from a PI regime to a more demanding PP regime (in terms of workability standards) may not improve social welfare in those cases in which a PI regime yields too much research and too little development. In order to emphasize the relevance of entry to providing an answer to the above question, it may be useful to establish the following

Proposition III.4 : If $\frac{d^3\gamma(v)}{dv^3}>0$, for any exogenously given number of entrants M a PI regime will always be welfare superior to a PP regime for $p<1$, with welfare levels under the two regimes converging at $p=1$.

As research expenditures, $M\varphi$, are obviously the same under the two regimes and social welfare is increasing in (expected) development expenditure, $E\{v\}$, it suffices to show that $E\{v^{PI}\}>E\{v^{PP}\}$.

Let $M=2$ (the extension to $M>2$ is straightforward); under a PP regime v^{PP} is chosen so as to

$$\max_v \; p\left[(1-p)\gamma(v,1)+p\gamma(v,2)-v\right]-\varphi \qquad \text{III.41}$$

where $\gamma(v,k)$ is gross profit (i.e. net profit plus development expenditure) in a k-firm oligopoly. The FOC for the maximization III.41 is

$$(1-p)\delta_1 g(v) + p\delta_2 g(v) = 1 \qquad \text{III.42}$$

where $\delta_k g(v)\equiv\frac{d\gamma(v,k)}{dv}$; let $g(v)$ be convex: $\frac{d^3\gamma(v)}{dv^3}>0$.

Let $\frac{1}{g(v)} \equiv \xi(v)$ (so that $\xi^{-1}(v)$ is convex), then:

$$v^{PP} \equiv \xi^{-1}\left((1-p)\delta_1+p\delta_2\right) \qquad \text{III.43}$$

Under a PI regime v^{PI} is the solution of the following:

$$\left[p(1-p)\max_v\{\gamma(v,1)-v\}+p^2\max_v\{\gamma(v,2)-v\}\right]-\varphi \qquad \text{III.44}$$

The FOCs for the above maximization are:

$\delta_1 g(v)=1$ i.e. $v=\xi^{-1}(\delta_1)$

$\delta_2 g(v)=1$ i.e. $v=\xi^{-1}(\delta_2)$

and thus finally:

$$E(v^{PI})\equiv p(1-p)\xi^{-1}(\delta_1)+p^2\xi^{-1}(\delta_2)>E(v^{PP})\equiv p\;\xi^{-1}\left((1-p)\delta_1+p\delta_2\right) \qquad \text{III.45}$$

where III.45 follows from the convexity of $\xi^{-1}(\cdot)$. ■

Of course, if there is free entry the two regimes cannot sustain the same number of entrants and the PP regime will always induce *fewer* firms to enter; indeed III.45 is also the condition that expected net profit under a PI regime be greater than under a PP regime.

By combining Propositions III.1 and III.3-4, we can now show that:

Proposition III.5: Provided discoveries are 'not too difficult' (in a sense made precise below), a switch from a PI regime to a PP regime is welfare-improving.

Suppose that $p=1$ so that the two regimes are in fact identical and at a free-entry equilibrium sustain the same number of firms (i.e. $N^{o}=n^{o}$) and that research costs, φ^{o} are such that $\pi^{PI}=\pi^{PP}=0$, with $N^{o}(=n^{o})$ being an integer. In view of Prop. III.3 we know that any change that brought about a decrease in entry would be welfare-improving; we can now show that the same result applies even under uncertainty. Suppose that the probability of success in research falls by a 'small' amount, i.e. $\hat{p}=1-\varepsilon$. Because of Prop. III.1, we know that at $\hat{p}$ $E\{\pi^{PI}\}>0$; let research costs $\hat{\varphi}$ be such to nullify the ('small') supernormal profits enjoyed by the N^{o} firms. Whereas these two small changes bring about correspondingly small changes under a PI regime, under a PP regime there will be discrete 'jumps'. In fact at $(\hat{p},\hat{\varphi})$ profits for the n^{o} $(=N^{o})$ firms are negative and thus a free-entry equilibrium will sustain $N^{o}-1$ firms (each earning supernormal profits). However, at $(\hat{p}, \hat{\varphi}, N^{\circ}-1)$ the net welfare levels under the two regimes are 'close' (recalling that, by Prop. III.4, they converge at $p=1$). Thus, taking into account the super-normal profits now generated by the $N^{\circ}-1$ firms, social welfare will be higher under a PP regime. ■

In order to ascertain how close to unity the probability of success in research has to be for the above argument to hold we can resort to the parametrization III.15-17; straightforward substitution yields

$$E\{W^{PI}\} \equiv \sum_{i=1}^{N}\binom{N}{i}p^{i}q^{N-i}\frac{i}{4\alpha}\, v_i^{PI} + N\, E\{\pi^{PI}\} \qquad \text{III. 46}$$

$$E\{\pi^{PI}\} \equiv \sum_{i=1}^{N}\binom{N}{i}p^{i}q^{N-i}\frac{1-2i\alpha}{2\alpha N}\, v_i^{PI} - \varphi \qquad \text{III. 47}$$

$$v_i^{PI} \equiv \left\{\frac{2i\alpha\theta^2}{(i+1)^2}\right\}^{1/(1-2\alpha)} \qquad \text{III. 48}$$

$$E\{W^{PP}\} \equiv \sum_{i=1}^{n}\binom{n}{i}p^{i}q^{n-i}\, \frac{i(i+2)}{2(i+1)^2}\, \theta^2 (v^{PP})^{2\alpha} - npv^{PP} - n\varphi \qquad \text{III. 49}$$

$$v^{PP} \equiv \left[\sum_{i=1}^{n}\binom{n}{i}\, \frac{i}{n}\, p^{i-1}q^{n-i}\, \frac{2i\alpha\theta^2}{(i+1)^2}\right]^{1/(1-2\alpha)} \qquad \text{III. 50}$$

Suppose that research costs φ are such that a free-entry equilibrium sustain exactly 3 firms under a PI regime; for any p we can now compute welfare levels for the PI regime and for the PP regime, bearing in mind that in the latter one fewer firm will be active.

Simulations suggest that the PP regime will yield a higher net expected social surplus for values of p around 0.5 (the switch-over value of p appear to be quite insensitive to changes in α, ranging from $p=0.506$ for $\alpha=0.01$ to $p=0.582$ for $\alpha=0.375$). When inventions are 'not too difficult' (i.e. $p>0.5$) and research cost are 'high' (so that at most 3 firms can be active) a switch from the entry-inducing PI regime to the concentration-inducing PP regime may in fact yield a welfare improvement. On reflection, this is in line with intuition: especially when p is "high" the two advantages of a PI regime — higher entry and higher ouput — may be more than offset by the advantages of a PP regime — lower industry research costs and more extensive development of research prototypes.

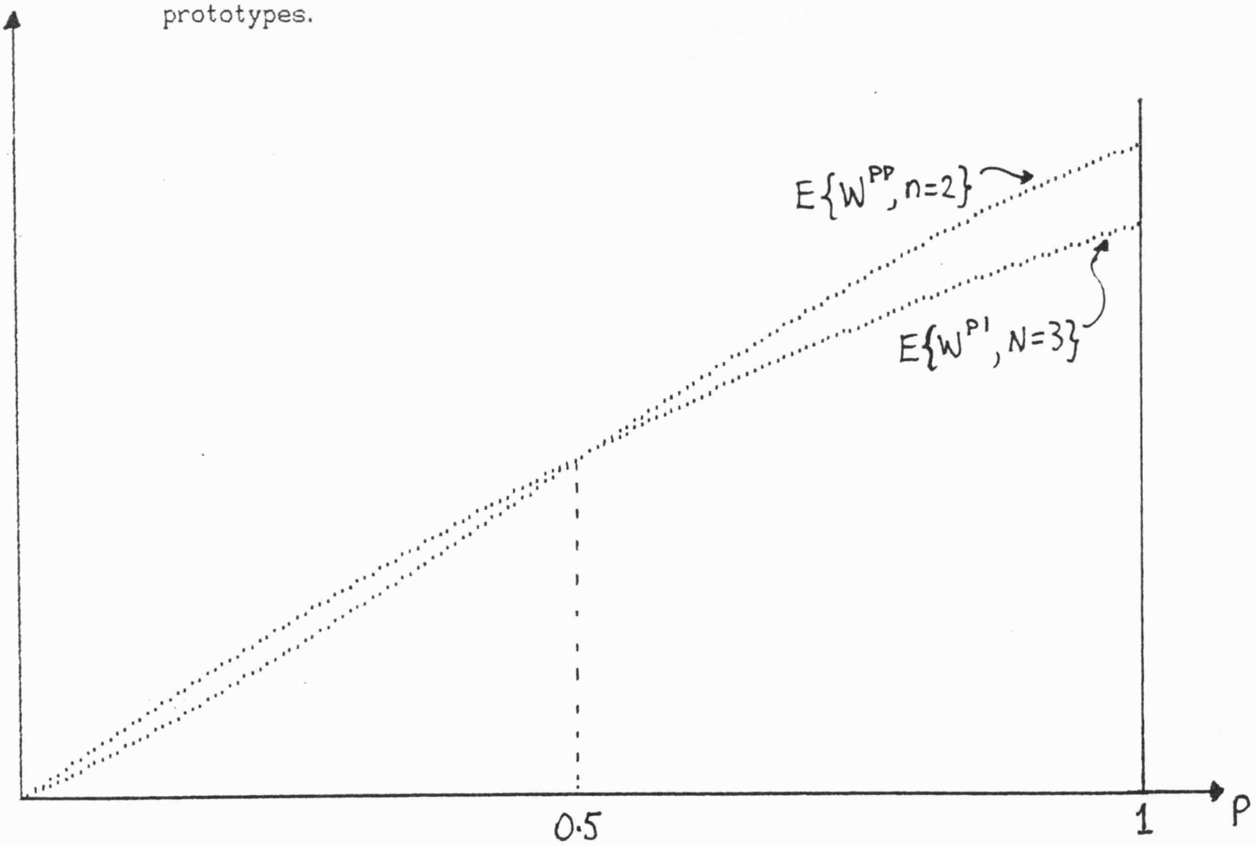


FIG. III.5

PI regime vs PP regime

III.8 Conclusions

In this chapter it has been suggested that if research and development are modelled as two separate stages, the patent system can take on the hitherto neglected and potentially significant role of affecting the allocation of resources between research and development. This is particularly important when the definition of product space is "permissive" so that multiple patents are awarded to close substitutes.

Of course, more research (and possibly more development) is required in the hitherto neglected area of the economics of patentability standards if economists are to meet the challenge posed by recent technological developments to the suitability of the current patent regime. As examples of industries to which the above R&D model could be profitably applied I shall consider the issue of the patentability of plant varieties and micro-organisms.

Consider the case of plant varieties first: as any variation in the set of features of a plant is considered a new variety, the current patent system defines the product space in a "permissive" fashion. Whether it is ideas- or product-based is less clear, in so far as current patentability standards demand that some development be undertaken (so as to achieve "stable" and "uniform" varieties), but on the other no development is required to satisfy the patent authorities that the "new" variety is not merely the result of "cosmetic breeding", but has valuable features in terms of yield, resistance to disease, etc. If it is plausibly assumed that the probability of succeeding in changing cosmetically the features of a plant variety is close to unity, then the above analysis suggests that it is likely that a switch to more stringent rules on the required development of varieties prior to patent filing could be welfare-improving.

A telling example of the economics of patents lagging behind the need for relevant policy advice is given by the European Commission's "Proposal for a Council Directive on the Legal Protection of Biotechnological Inventions" (1988). The proposed Directive correctly singles out a properly designed patent system as an important, indeed as a crucial, factor for the success of the European biotechnology industry, but is forced by the lack of relevant economic models of R&D to ground its recommendations on 'common sense' (i.e. on unrigorous economic theorizing). Indeed in the document one can find various examples of the implicit assumption that encouraging competition at the research stage is always more socially beneficial than favouring the wider dissemination of new ideas — an assumption that the above analysis has shown to be valid only for certain ranges of the parameters that describe technology and demand.

CHAPTER IV

R&D, Quality, and the Integer Constraint

IV.1 Introduction

As we have seen in the preceding chapters, the 'integer constraint' due to the indivisibility of firms brought about by the 'lumpy' nature of expenditures in R&D is not just a technical nuisance that can be dismissed with the popular proviso "assuming the number of firms can be treated as a continuous variable ... ", but often has far-reaching consequences in so far as it may **reverse** the welfare implications derived from a non-integer-constrained version of the model.

In this chapter, we shall analyse in some detail the economics of the integer constraint and shall suggest that its shadow looms larger than is commonly recognized. As a by-product of the analysis we shall also show that two major strands of the economics of technical change, namely the economics of quality and the economics of R&D, are not as independent as the literature might suggest. Finally, it shall be argued that even in a simple Cournot-Nash oligopoly model with no R&D a proper formalization of the integer constraint can be useful to account, for example, for the positive correlation between profitability and concentration in a model **with free entry**.

IV.2 The Missing R&D-Quality Link

In the 1970s a large literature[1] analysed the topic of **quality competition** and in particular the effects of given market structures on the quality of output, as measured by a scalar, q. In the 1980s the emphasis moved on the one hand on modelling market structure in vertically differentiated industries, and on the other on modelling R&D within an explicitly game-theoretic framework.

In this section, I shall argue that interesting insights can be gained by establishing a link between quality-competition models and game-theoretic models of R&D. To this end I shall interpret Dasgupta and Stiglitz (1980) model of R&D in the light of quality competition theory.

It shall be recalled that Dasgupta and Stiglitz explore, among other things, the R&D behaviour in a socially-managed industry and in a monopoly. Let the social planner's and the monopolist's payoffs be, respectively:

$$U(Q) - C(x, Q) \qquad \text{IV. 1}$$

$$p(Q)Q - C(x, Q) \qquad \text{IV. 2}$$

where $U(Q)$ is gross social welfare, Q is output, x is expenditure in R&D, $p(Q)$ is the demand function and $C(x, Q)$ is a cost function that Dasgupta and Stiglitz specialise to:

$$C(x, Q) = c(x)Q + x \qquad \text{IV. 3}$$

---

[1] For a survey, see Schmalensee (1979).

By using specific functional forms for $U(Q)$ and $c(x)$, Dasgupta and Stiglitz show a socially managed industry will engage in more substantial cost reductions than a monopoly, i.e. $x^S > x^M$.

It is worth considering the cost function $C(x, Q)$ in some detail; whereas, for a given x^0, $c(x^0)Q$ can be interpreted as a proper minimum cost function, i.e. derived from $\{\min_{z \geq 0} w.z$ s. t. $Q=Q(z)\}$ where z = column vector of inputs, w = row vector of input prices, $Q(z)$ = linear homogeneous production function, $C(x, Q)$ is a 'hybrid' cost function, whose arguments are output Q and an input, x.

Simply by recasting $C(x, Q)$ in a more conventional way we can establish the correspondence between Dasgupta and Stiglitz's typical model of R&D with a typical quality-competition model. This re-interpretation has several advantages: on the one hand it shows that the result $x^S > x^M$ that Dasgupta and Stiglitz have proved for specific functional forms for $U(Q)$ and $c(x)$, holds for more general cases, but, on the other hand, it highlights the fact the this result depends on a specific assumption regarding the quality-improving technology.

Indeed, the missing link between Dasgupta and Stiglitz's R&D model and a typical quality-competition model is furnished by a quality-improving technology. Whereas the latter model does not specify how quality improvements are brought about (presumably by investing in R&D), Dasgupta and Stiglitz do not spell out how

R&D expenditure can reduce costs of production (presumably by a qualitative change in the industry's capital equipment).

Let $q=I(x)$ be the quality-improving function that maps expenditure in R&D, x, on to a scalar quality index q, with $I_x(x)>0$.

Defining $c(I^{-1}(q))\equiv G(q)$, $C(x,Q)$ can be written as

$$\tilde{C}(q,Q) = G(q)Q + I^{-1}(q) \qquad \text{IV. 4}$$

At this stage we can profitably resort to a standard result in quality competition theory[2] which states that if $p_{qQ}\geq 0$ and $\tilde{C}_{qQ}<0$, then a monopolist will produce goods of inferior quality as compared with a socially managed industry, i.e. $q^S>q^M$. It is easy to see that both assumptions hold for the special case examined by Dasgupta and Stiglitz, in which R&D does not affect the demand function ($p_{qQ}=0$) and $\tilde{C}_{qQ}\equiv G_q(q)\equiv\frac{c_x(x)}{I_x(x)} < 0$ and thus, as $I_x(x)>0$, it follows that $\{q^S>q^M\}\Leftrightarrow\{x^S>x^M\}$, which is precisely the result obtained by Dasgupta and Stiglitz.

Thus by highlighting the link between R&D and quality[3] we have

---

[2] See, for example, Sheshinski (1976) and below, sec. IV.3.

[3] Consider the following quotation from Dasgupta and Heal (1979), p. 475:
The problems of modelling the effects of R and D are … particularly acute: it is necessary both to describe the occurence of technological change and to specify the relationship between the allocation of effort to R and D and the resulting effect on changes in technology. (…) it is fortunate that a number of important points can be made without any explicit modelling of the R and D - technological change link.
What we are arguing in this section is that even a very modestly explicit modelling of the link between R&D and quality can shed light on the nature and limitations of the underlying model of R&D.

shown that:

(i) on the one hand the result that a socially managed industry will invest a larger amount in R&D is not confined to the specific functional forms used by Dasgupta and Stiglitz;

(ii) on the other hand, the above result applies to specific relationships between demand and quality (i.e. $p_{qQ} \geq 0$) and marginal cost and quality (i.e., $\tilde{C}_{qQ} < 0$).

As a final remark on the correspondence between R&D models and models of quality competition, it may be noted that by exploring the quality-improvement implications of Dasgupta & Stiglitz's model, the latter can be interpreted in novel ways. Consider the large class of cases in which there exists a quantifiable difference between **physical** and **vendible** output. For example, the quality of output can be stochastic and for some reason (whether legal or strategic) firms cannot - or do not wish to - sell output below a given minimal standard. Thus output must go through a quality inspection process that transforms physical output **Y** into vendible output **Q**. The efficiency of the quality-check process depends on R&D expenditure: let **e(x)** be the percentage of output that can be marketed if **x** units of R&D are expended, i.e.

$$Q = e(x)Y; \quad e_x(x) > 0 \qquad \text{IV.5}$$

As welfare and market price depend on vendible, rather than physical output, the planner's and the monopolist's payoffs will be, respectively

$$U(Q) - c \frac{Q}{e(x)} - x \qquad \text{IV.6}$$

$$P(Q)Q - c \frac{Q}{e(x)} - x \qquad \text{IV.7}$$

which, apart from a change in notation, are identical to IV.1-3.

IV.3 Product quality and the integer constraint[4]

Having shown the correspondence between R&D models and models of quality competition, we can now examine the effects of the integer constraint in a model of quality competition with an endogenously determined market structure.

All models that determine the equilibrium number of firms endogenously by treating it as a **continuous** variable acknowledge that such a patently counterfactual assumption is made for convenience and on the belief that allowing for the integer constraint would leave the qualitative results of the model unchanged.

The purpose of this section is to show by means of a simple example that in the context of a product-quality model the integer constraint does make a substantial qualitative difference[5]. Whereas by treating the number of firms as a

---

[4] A shortened version of this section has been published in La Manna (1987a).

[5] Although the emphasis in this section is on product quality, it should be borne in mind that the key issue of the relevance of the integer constraint applies to a far larger class of models - indeed to any model that treats market structure as endogenous.

continuous variable one obtains the result that in a free-entry oligopoly there is **excess entry** and that the equilibrium level of quality may **differ** from the socially optimal one, when the integer constraint is allowed for, the socially optimal (integer-constrained) number of firms may **exceed** the (integer-constrained) number of firms sustained in oligopoly and product quality may be **the same** under the two regimes.

In sec. IV.3.1 the standard product quality model with be extended to the case of a free-entry oligopoly; sec. IV.3.2 contains an example demonstrating the the qualitative results derived in IV.3.1 do not necessarily carry over to the integer-constrained solution. All proofs and technical details are relegated to Appendix IV.

IV.3.1 Product quality under a socially-managed industry and free entry oligopoly

Some preliminary remarks on the literature are in order.

Almost all of the early models of product quality competition analysed the effects of a **given** market structure on product quality, as measured by a scalar, **q**. The standard result was that **q** is independent of the quantity of output produced, **Q** (and thus of market structure) iff **both** the following conditions are satisfied:

$$P_{qQ}(q, Q) = 0 \qquad \text{IV.8}$$

$$C_q(q, Q) - QC_{qQ}(q, Q) = 0 \qquad \text{IV.9}$$

where $P(\cdot)$ and $C(\cdot)$ are the inverse demand function and the cost function, respectively.

The theory of product quality underlying the above result is quite different from the theory of vertical product differentiation pioneered by Jaskold Gabszewicz and Thisse (1979, 1980) and Shaked and Sutton (1982, 1983) in so far as the latter analyse the case in which consumers are heterogenous (i.e. attach different values to quality and vary in their income endowments) and, more importantly, consider market structure as endogenous.

The model developed below can be viewed as an intermediate step between the 'old' and the 'new' theory of product quality, in that, like the latter, it aims at determining market structure but, like the former, takes as its starting point a very simplified notion of product quality.

The reason for not considering the more interesting case of heterogenous consumers is that the simpler case of an 'add-on' quality premium valued equally by all (identical) consumers affords the most economical way of making the point on the relevance of the integer constraint.

Thus the demand function is assumed to be additively separable in output Q and product quality q:

$$P(q, Q) = f(Q) + g(q) \qquad \text{IV. 10}$$

with $g(0)=0$ and $P(q, 0)=0$. That is to say, the quality-augmented demand curve $P(q, Q)$ is simply the demand curve for the minimal-quality good ($q=0$) shifted to the right by the amount of value of quality premium $g(q)$. Of course, the functional form IV.10 guarantees that $P_{qQ}(q, Q)=0$.

Turning now to the cost side, special care must be taken in modelling quality-related costs. In fact, in so far as quality is obtained through expenditures in R&D, it is bound to emerge as a fixed cost in $C(q, Q)$. And yet the early literature on product quality abounds of examples that assume away the fixed cost feature of quality change. Take, for instance, the following functional form, used by Levhari and Peles (1973) and Schmalensee (1979), which rules out any quality-related fixed cost:

$$C(q, Q) = H(Q) + QG(q) \qquad \text{IV. 11}$$

Obviously without some sort of fixed cost one cannot determine the equilibrium number of firms and *a fortiori* consider the effects of the integer constraint. We can then introduce quality-related fixed costs explicitly:

$$C(q_i, Q_i) = F(q_i) + V(q_i, Q_i) \qquad \text{IV. 12}$$

Notice that IV.12 assumes that there are no spill-over effects (costs for a typical firm i depend only on its own q_i and Q_i).

Output quality and quantity in a socially managed industry.

Assuming complete information, a planner can locate her first-best optimum by choosing the combination of Q^s (output per plant), q^s (quality index) and n^s (number of plants) that maximizes net surplus, i.e.

$$W(q^s, Q^s, n^s) \equiv \max_{q,Q,n} \int_0^{nQ} [f(v)+g(q)]dv - n[F(q)+V(q,Q)] \qquad \text{IV. 13}$$

i.e. q^s, Q^s, n^s are the solution of the FOCs:

$$W_Q(q,Q,n)=n[f(n^sQ^s)+g(q^s)-V_Q(q^s,Q^s)]=0 \qquad \text{IV. 14}$$

$$W_q(q,Q,n)=n[Q^s g_q(q^s)-F_q(q^s)-V_q(q^s,Q^s)]=0 \qquad \text{IV. 15}$$

$$W_n(q,Q,n)=Q[f(n^sQ^s)+g(q^s)]-F(q^s)-V(q^s,Q^s)=0 \qquad \text{IV. 16}$$

Notice that IV.14 and IV.16 yield the unsurprising result that, for any given q, $MC(q)=AC(q)=P(q)$, while IV.15 says that, for any given Q, product quality is chosen so as to equate its marginal benefit, $Qg_q(q)$, to its marginal cost, $F_q+V_q(q,Q)$.

Output quality and quantity in a free-entry oligopoly

All firms are assumed: (i) to have access to the same technology; (ii) to behave non-cooperatively; (iii) to entertain Cournot conjectures regarding their control variables, q and Q; (iv) to determine q and Q simultaneously. These assumptions guarantee that the resulting Cournot-Nash equilibrium is symmetric. Furthermore, it is assumed that there are no barriers to entry and that entry take place as long as entrants make non-negative profits.

The additivity of $P(q,Q)$ in q and Q simplifies greatly the modelling of this oligopoly environment in so far as it rules out any product selection problem: goods produced by all oligopolists are perfect substitutes and are sold at unit prices equal to a quantity-determined common constant plus a quality-related premium. Defining

$$\pi(q^o, Q^o) \equiv \max_{q_i, Q_i} \{ f(Q_{-i}+Q_i) + g(q_i)Q_i - F(q_i) - V(q_i, Q_i) \} \qquad \text{IV. 17}$$

and recalling that, at a free-entry equilibrium $\pi(q^o, Q^o, n^o)=0$ the solution (q^o, Q^o, n^o) is defined by:

$$\pi_Q(q, Q) = f(n^o Q^o) + Q f_Q(n^o Q^o) + g(q^o) - V_Q(q^o, Q^o) = 0 \qquad \text{IV. 18}$$

$$\pi_q(q, Q) = Q^o g_q(q^o) - F_q(q^o) - V_q(q^o, Q^o) = 0 \qquad \text{IV. 19}$$

$$\pi(q, Q) = Q^o f(n^o Q^o) + Q^o g(q^o) - F(q^o) - V(q^o, Q^o) = 0 \qquad \text{IV. 20}$$

Comparison between Social Optimum and Oligopoly Equilibrium

In the Appendix to this Chapter it is shown that:

$$n^o > n^s \qquad \text{IV. 21}$$

$$Q^o < Q^s \qquad \text{IV. 22}$$

$$q^o \gtrless q^s \qquad \text{IV. 23}$$

according as to whether $F_q(q) + V_q(q, Q) \lessgtr Q V_{qQ}(q, Q)$.

These results are simply a straightforward generalization of standard product-quality models that take n as exogenously given (see, for example, Sheshinski (1976)).

IV.3.2 The importance of the integer constraint

Because of quality-related fixed costs, $F(q)$, firms are non-divisible and hence n cannot be treated as a continuous variable. Indeed, as the probability of n being an integer is a set of measure zero, it is 'almost certain' that n^o will not be an integer and therefore (n^o, q^o, Q^o) cannot be a Cournot-Nash equilibrium.

A simple way out of this *impasse* is afforded by recasting the model as a two-stage game: the first stage is just as described above, whereas in the second stage $|n^o|$ firms, where $|n^o|$ is the largest integer not exceeding n^o, re-compute the Cournot-Nash equilibrium of the model, conditional on there being $|n^o|$ entrants.

The key point to notice is that whereas $|n^o|$ cannot exceed n^o (for profits would be negative for each of the $(|n^o|+1)$ entrants), no such uni-directional restriction applies to a social planner, who, faced with a non-integer n^s, can adjust the number of firms either downwards or upwards. In the latter case, firms, of course, would be operating on the falling segment of their marginal cost, but the loss due to marginal-cost pricing would be more than offset by the larger consumers' surplus.

As a result of this asymmetry, it is conceivable that, even though in the integer-unconstrained case there is always excessive entry, i.e. $n^o>n^s$, once the integer constraint is taken into account, the inequality may be reversed, i.e. $|n^o|<|n^s|$, where $|n^s|$ is the integer-constrained value of n^s.

Of course, the integer constraint may yield inequality reversals regarding the number of firms with any sort of fixed costs (not necessarily quality related), but, as the following example shows, when fixed costs *are* quality-related, such reversals change the qualitative results derived from the integer-unconstrained model.

IV.3.3 An Example

Suppose that $P(q, Q)$ is linear in output, i.e.

$$P(q, Q) = a - \sum Q + g(q) \qquad \text{IV. 24}$$

Substituting IV.24 into the FOCs that determine the social optimum and the free-entry oligopoly equilibrium we obtain:

$$W_Q(q, Q, n)=n\left[a-n^s Q^s+g(q^s)-V_Q(q^s, Q^s)\right]=0 \qquad \text{IV. 25}$$

$$W_q(q, Q, n)=n\left[Q^s g_q(q^s)-F_q(q^s)-V_q(q^s, Q^s)\right]=0 \qquad \text{IV. 26}$$

$$W_n(q, Q, n)=Q\left[a-n^s Q^s+g(q^s)\right]-F(q^s)-V(q^s, Q^s)=0 \qquad \text{IV. 27}$$

$$\pi_Q(q, Q)=a-(n^o+1)Q^o+g(q^o)-V_Q(q^o, Q^o)=0 \qquad \text{IV. 28}$$

$$\pi_q(q, Q)=Q^o g_q(q^o)-F_q(q^o)-V_q(q^o, Q^o)=0 \qquad \text{IV. 29}$$

$$\pi(q, Q, n)=Q^o[a-n^o Q^o+g(q^o)]-F(q^o)-V(q^o, Q^o)=0 \qquad \text{IV. 30}$$

Comparing IV.25-27 with IV.28-30, it is easy to see that a sufficient condition for the two regimes to yield the same output per firm and the same product quality level, irrespectively of the sign of $F_q(q)+V_q(q, Q)-QV_{qQ}(q, Q)$, is that

$$|n^s| = (|n^o|+1) \qquad \text{IV. 31}$$

In fact, under this condition, the FOCs with respect to q and Q would be identical for the two regimes, and IV.27 and IV.30 do not apply in the second stage of the game, when n^s and n^o are integer- constrained and exogenous.

Parametrizing $P(q,Q)$ and $C(q,Q)$ as follows:

$$P(q,Q) = 51.7 - \sum Q - \frac{100}{q} \qquad \text{IV.32}$$

$$C(q,Q) = q^2 + 1.5qQ^2 \qquad \text{IV.33}$$

routine calculations confirm that IV.21-23 hold:

$$Q^s = 1.58 > Q^o \simeq 1.44$$

$$q^s = 3.75 > q^o \simeq 3.7$$

$$n^s = 4.58 < n^o \simeq 4.96$$

Let $(|q^o|, |Q^o|)$ and $(|q^s|, |Q^s|)$ be, respectively, the profit- and welfare-maximizing levels of product quality and output per firm when the number of firms is integer-constrained (i.e. the solutions of the second stage of the game). It is easy to confirm that a social planner would attain a higher level of net social welfare by adjusting n^s **upwards** (as shown in Fig. IV.1), i.e.

$$W(|q^s|, |Q^s|, |n^s|+1) > W(|q^s|, |Q^s|, |n^s|)$$

Therefore, as condition [IV.31] holds, i.e. $|n^s| = (|n^o|+1)$, we obtain

$$|n^s| = 5 > |n^o| = 4$$

$$|Q^s| = |Q^o| \simeq 1.53$$

$$|q^s| = |q^o| \simeq 3.73$$

Notice that contrary to the implication of the integer-unconstrained case, product quality is the same under the two regimes, even though $F_q(q)+V_q(q,Q) > QV_{qQ}(q,Q)$.

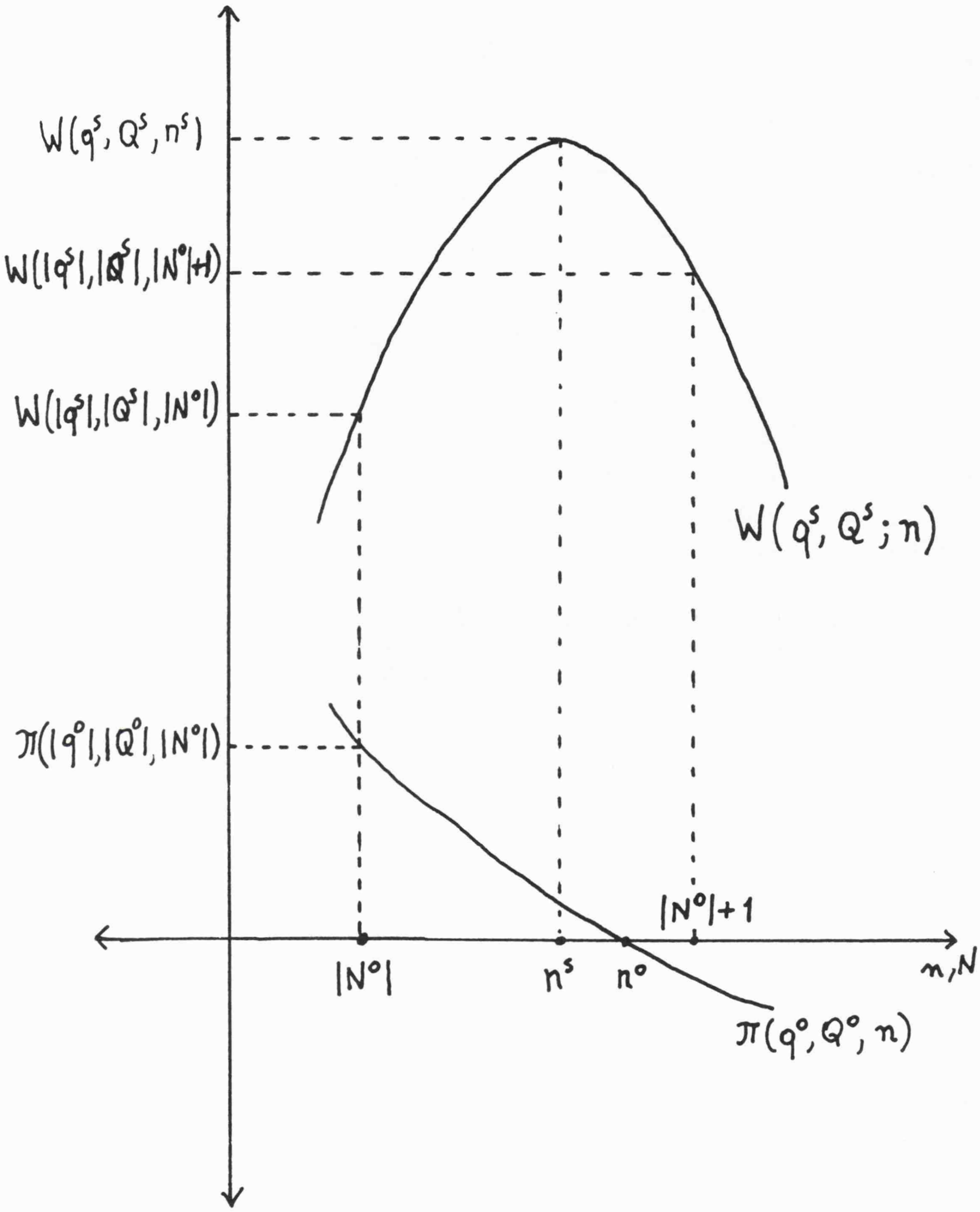


Fig. IV. 1

The Importance of the Integer Constraint

IV.4 Profits, Concentration, and the Integer Constraint[6]

In the example sketched in the previous section each of the $|n^0|$ oligopolist enjoys strictly positive super-normal profits, in spite of the industry being modelled as a free-entry oligopoly. This shows quite clearly that as soon as the integer constraint is taken into account, free entry and positive profits are no longer incompatible. The question can then be raised as to whether the integer constraint can account for the positive correlation between profits and concentration under free-entry conditions.

Prima facie there is no mystery in the commonly observed empirical correlation between concentration and profitability. The answers that most industrial economists would provide if asked to provide a theoretical explanation for it would point to the fact that (i) in a cooperative model, high concentration leads inevitably to higher profitability because of increased collusion – the fewer firms there are, the easier collusion is, thereby raising individual and joint profits; (ii) in a non-cooperative model, as long as concentration is defined as an inverse function of **N**, the number of firms in the industry, then, provided firms choose output levels simultaneously, it can be shown [see Seade (1980)] that profits per firm fall as **N**

---

[6] This section draws on La Manna (1986).

rises.

Neither of the above explanations refers explicitly to entry conditions – and for a good reason: if it is assumed that there are no barriers to entry, then, in so far as free entry implies **zero** super-normal profits **irrespectively** of the value of **n**, there ought to be no correlation at all between profitability and concentration.

It follows that, according to the received wisdom, no theoretical explanation of the profitability-concentration link is available if one makes the reasonable joint assumption of free entry and non-cooperative behaviour[7].

In this section it will be shown that because of the integer constraint arising from the indivisibility of firms, concentration and profitability are correlated even in a non-cooperative oligopoly with free entry. It will be shown that super-normal profits depend on two structural parameters:

1. the size of entry fees; the resulting **'size effect'** is unambiguously positive in the sense that, *ceteris paribus*, industry with high entry fees are characterized by high concentration and high profits per firms;
2. the extent to which the integer constraint is binding; the sign of the resulting **'integer constraint effect'** is ambiguous and in some special albeit unlikely cases may even offset the 'size effect' and produce a perverse **negative**

[7] For an interesting model with both free entry and collusion, see Brander and Spencer (1985).

correlation between concentration and profitability.

Finally, it will be shown that, provided the distributions of entry fees across industries are 'well-behaved' (in a sense made precise below), then cross-industry regressions will show, on average, a positive correlation between concentration and profitability.

The logic of the argument is simple and can be illustrated as follows: suppose that the number of firms sustained by a free-entry equilibrium in industry i (respectively, j) is 2.5 (4.5). Obviously only two (four) firms can be active, each earning super-normal profits. One would expect that, *ceteris paribus*, the more concentrated industry i to be more profitable than industry j. This is because, figuratively speaking, the spoils of half a firm are worth more in industry i and, moreover, are shared among fewer firms. However, it is not difficult to see that, because of the integer constraint, more concentrated industries may turn out to be *less* profitable. Let n_i and n_j be respectively 2.01 and 3.99; in industry i profits will be close to zero, whilst the three firms active in industry j will share the non-negligible profits that would have accrued to the "ninety-nine hundredths of a firm" which would have existed, had not been for the integer constraint.

What follows formalizes the intuitive argument sketched above.

IV.3.1 Free-entry Cournot Equilibrium and the Integer Constraint

Consider an economy with **M** free-entry industries indexed by j (j=1, . . , **M**). Assume (for simplicity only) that each industry is faced with a linear inverse demand function for its homogenous product and that each firm has constant marginal costs, i.e.

$$P_j = a_j - b_j \sum_{j}^{n_j} x_{ij} \qquad i=1,\ldots,n_j \qquad \text{IV. 34}$$

$$C_{ij}(x_i) = c_j x_{ij} + F_j \qquad \text{IV. 35}$$

where: x_{ij}=ith firm's output level in industry j

n_j=number of firms in industry j

F_j=industry-specific fixed entry fee in industry j (on a per-period basis).

In order to focus on inter-industrial differences in entry fees, it may be assumed that all industries are identical in all respects but entry fees (nothing of substance hinges on this notation-saving assumption) :

$$a_j = a; \quad b_j = b; \quad c_j = c \qquad \forall j \qquad \text{IV. 36}$$

Notice that, as firms are identical, the reciprocal of the number of firms can be used as an unambiguous index of concentration. Given the number of firms in industry j, a Cournot-Nash equilibrium is established. Straightforward calculations show profits per firm to be

$$\Pi_j = (a-c)^2/b(n_j+1)^2 - F_j \qquad \text{IV. 37}$$

The free-entry equilibrium number of firms in industry j,

n_j^{**}, is determined by the following zero-profits condition:

$$F_j = (a-c)^2/b(n_j^{**}+1)^2 \qquad \text{IV.38}$$

Of course, the probability of n_j^{**} being an integer is a set of measure zero; thus, let $|n_j|$ be the integer-constrained number of firms in industry j, i.e., the largest integer not exceeding n_j^{**}.

It can be seen that industrial structure is modelled as a perfect equilibrium of a two-stage game: in the first stage firms simultaneously decide whether to incur the non-recoverable fixed entry fee F_j. In the second stage, those firms (numbering n_j^{**}) that have paid F_j choose their profit-maximizing output levels.

Actual profits per firm, $|\Pi_j|$, are given by

$$|\Pi_j| = (a-c)^2/b(|n_j|+1)^2 - (a-c)^2/b(n_j^{**}+1)^2 \qquad \text{IV.39}$$

i.e., defining $N_j \equiv |n_j|+1$ and $n_j^{**}-|n_j| \equiv \mu \in (0, 1)$, we obtain

$$|\Pi_j| = (a-c)^2(2N_j+\mu_j)\mu_j/bN_j^2[N_j^2+(2N_j+\mu_j)\mu_j] \qquad \text{IV.40}$$

It can be easily verified that, for any $\mu_i \leq \mu_k$,

$$|\Pi_i| \underset{<}{\overset{>}{=}} |\Pi_k| \iff |n_i| \underset{>}{\overset{<}{=}} |n_k| \qquad \text{IV.41}$$

According to IV.41, concentration – as measured by $1/|n_j|$ – and profitability – as measured by $|\Pi_j|$ – are **positively** correlated. However, from IV.40 it can also be seen that actual profits depend not only on the **size** of fixed entry fees ('**size** effect'), but also on the difference between the free-entry equilibrium of firms with and without fractional entry, i.e. on μ (**'integer constraint effect'**).

Whilst the size effect always leads to a positive correlation between concentration and profits (see IV.41), the integer constraint effect can work in the opposite direction; indeed, cases may arise in which the latter more than offsets the former, giving rise to an overall **negative** correlation between concentration and profitability[8].

In order to show that, on average, the size effect can be expected to be stronger than the integer constrain effect, thereby accounting for the observed positive correlation between profits and concentration, the following facts should be taken into account:

Fact 1. Industry j will sustain $|n_j|$ firms at a free-entry equilibrium iff

$$F_j \epsilon \left[\frac{(a-c)^2}{b(|n_j|+2)^2}, \frac{(a-c)^2}{b(|n_j|+1)^2}\right) \equiv [F_j^{min}, F_j^{max})$$

Fact 2. The length of the above half-open interval falls with $|n_j|$.

Fact 3. For any given $|n_j|$, $|\Pi_j|$ is an increasing concave function of μ_j, as shown in Fig. IV.2(a).

Let J be the set of industries with $|n_j|$ active firms and let $\delta_j(F_j)$ be the distribution of fixed entry fees over

---

[8] Consider the following example: $a=2$, $b=c=1$, $F_1=2.\ 1^{-2}$ $F_2=3.\ 9^{-2}$. From IV.38 it follows that $|n_1|=1$, $|n_2|=2$, $\mu_1=0.\ 1$, $\mu_2=0.\ 9$. Using IV.39 we can compute actual (super-normal) profits : $|\Pi_1|\simeq 0.0232 < |\Pi_2|\simeq 0.0456$.

$[F_j^{min}, F_j^{max})$; needless to say there is a one-to-one correspondence between $\delta_j(F_j)$ and a suitable distribution $\sigma_j(\mu_j)$ defined over $(0, 1)$.

In order for a cross-industry regression of **average** profits onto any index of concentration based on the number of active firms to yield a **negative** coefficient, the distributions $\delta_j(F_j)$ must have the unusual property that the difference between F_j^{max} and $\hat{F}_j$ (the average fixed entry fee) is a decreasing convex function of $|n_j|$. Or, to put it differently, the average 'integer constraint effect' (i.e. the average value of μ_j) must be **larger** in **less** concentrated industries.

In the absence of any economic reason why the **average size** of fixed entry fees should be systematically related to the **difference** between maximum and average fixed entry fees (i.e. $F_j^{max} - \hat{F}_j$), we must conclude that the **size effect** can be expected, on average, to be stronger than the **integer constraint effect**, thereby yielding a **positive** correlation between concentration and profitability.

In conclusion, in this section has shown that profitability and concentration are correlated even in a model with **both** free entry **and** no collusion. This is because with positive entry fees and non-fractional entry, free entry is compatible with positive super-normal profits being earned by active firms. These profits, however, are not correlated with concentration as such, but rather with the size of entry fees and the difference between the free-entry equilibrium of firms with and without fractional entry, i.e. μ_j. For a positive correlation to exist, it is sufficient that, on average, the latter is not inversely related to the size of fixed entry fees.

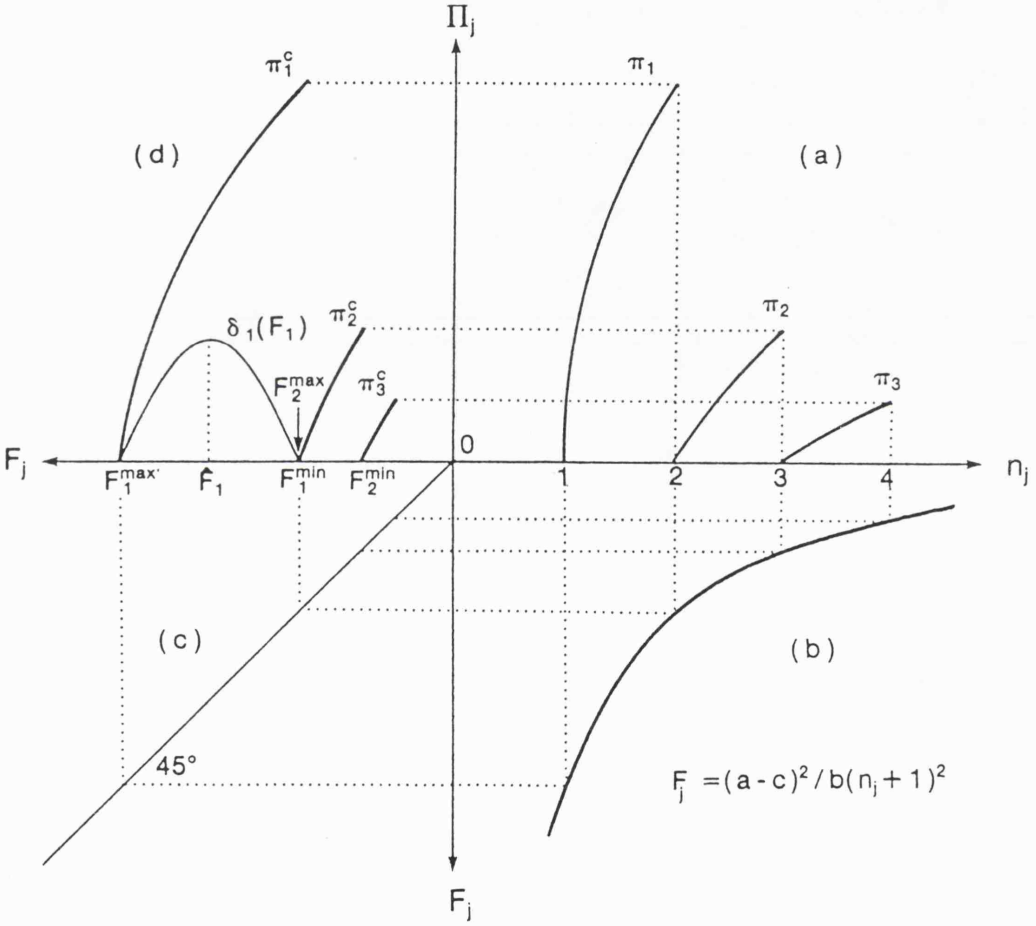


Fig. IV. 2

Size Effect vs Integer Constraint Effect

APPENDIX IV.A

The FOCs IV.14 and IV.16 entail that for any given product quality level, q, $AV(q)=MC(q)$, i.e.:

$$\mu(q,Q)\equiv \frac{F(q)+V(q,Q)}{Q} - V_Q(q,Q)=0 \qquad \text{IV.A.1}$$

Totally differentiating IVA.1 and IV.15 we obtain respectively:

$$\left.\frac{dq}{dQ}\right|_{\mu(q,Q)=0} = \frac{QV_{QQ}(q,Q)}{F_q(q)+V_q(q,Q)-QV_{qQ}(q,Q)} \qquad \text{IV.A.2}$$

$$\left.\frac{dq}{dQ}\right|_{W_q(q,Q)=0} = \frac{g_q(q)-V_{qQ}(q,Q)}{g_{qq}(q)+F_{qq}(q)+V_{qq}(q,Q)} \qquad \text{IV.A.3}$$

From the SOCs for the maximization of $W(q,Q,n)$ - which are assumed to be satisfied - it can be seen that the slope of the curve $q=\tilde{q}(Q)$ defined implicitly by $\mu(q,Q)=0$ is steeper than the slope of the curve $q=\hat{q}(Q)$ defined by $W_q(q,Q)=0$ and that the sign of the slope of both curves equals the sign of $F_q(q)+V_q(q,Q)-QV_{qQ}(q,Q)$; see Fig. IV.A.1 and IV.A.2.

Total differentiation of IV.14 and IV.16 for any given q yields

$$\left.\frac{dn}{dQ}\right|_{W_Q(q,Q,n)=0} = - \frac{n\left(nf_Q(nQ)-V_{QQ}(q,Q)\right)}{nQf_n(nQ)+f(nQ)+g(q)+V_Q(q,Q)}<0 \qquad \text{IV.A.4}$$

$$\left.\frac{dn}{dQ}\right|_{W_n(q,Q,n)=0} = - \frac{nQf_Q(nQ)+f(nQ)+g(q)-V_Q(q,Q)}{Q^2f_n(nQ)}<0 \qquad \text{IV.A.5}$$

For any given q, $W_Q(q,Q,n)=0$ yields a curve $n=\tilde{n}(Q)$ that is steeper than the curve $n=\hat{n}(Q)$ implied by $W_n(q,Q,n)=0$; the SOCs guarantee that both curves are downward sloping (see Fig. IV.A.3).

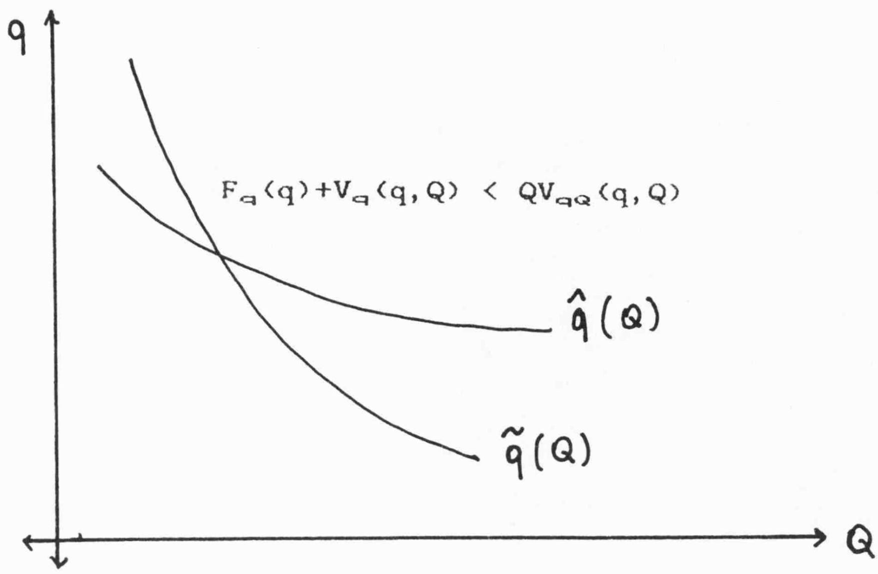


Fig. IV.A.1

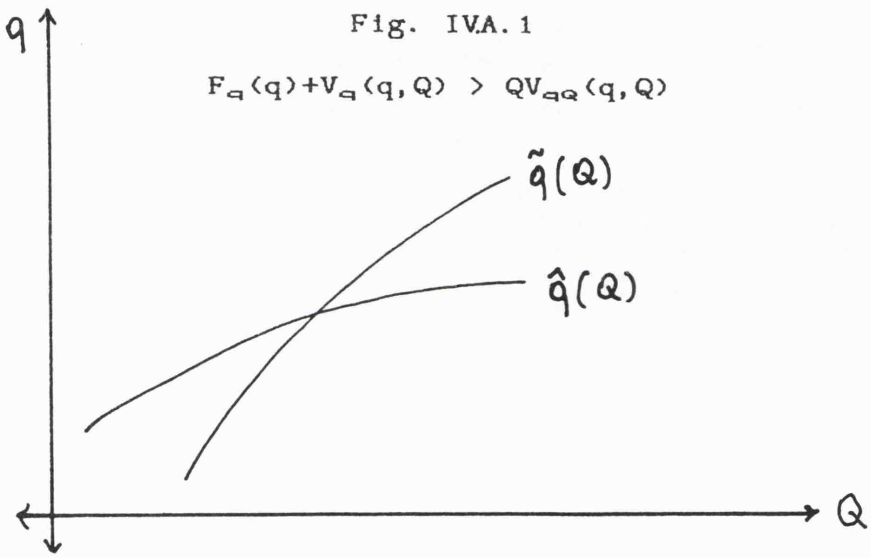


Fig. IV.A.2

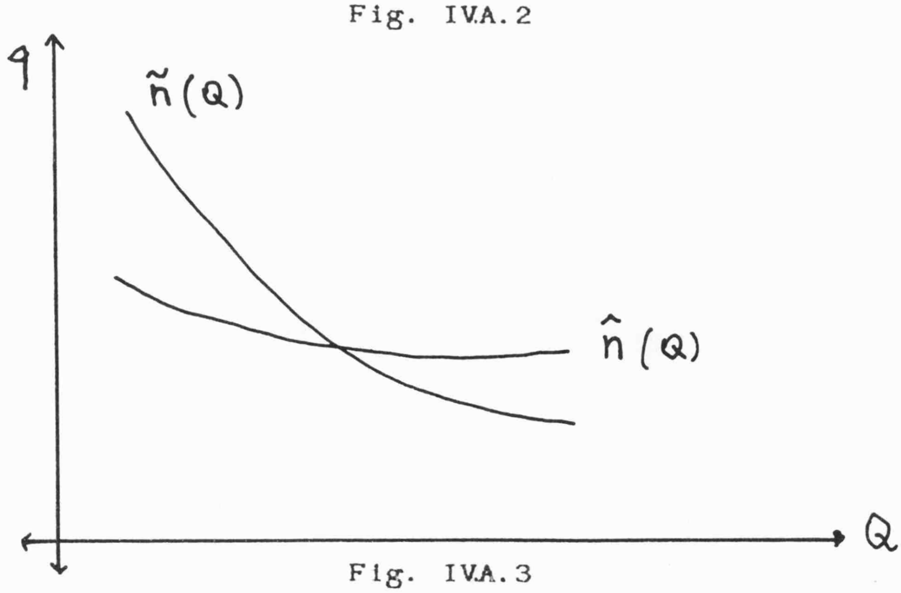


Fig. IV.A.3

Turning now to the free-entry oligopoly case, by combining the two FOCs IV.18 and IV.20, we obtain:

$$\lambda(q,Q) \equiv \frac{F(q)+V(q,Q)}{Q} - V_Q(q,Q)+Qf_Q(nQ)=0 \qquad \text{IV.A.6}$$

Total differentiation of IV.A.6 and IV.19 yields

$$\left.\frac{dq}{dQ}\right|_{\lambda(q,Q)=0} = -\frac{2f_Q(nQ)+Qf_{QQ}(nx)-QV_{QQ}(q,Q)}{F_q(q)+V_q(q,Q)-QV_{qQ}(q,Q)} \qquad \text{IV.A.7}$$

$$\left.\frac{dq}{dQ}\right|_{\pi_q(q,Q)=0} = -\frac{g_q(q)-V_{qQ}(q,Q)}{g_{qq}(q)+F_{qq}(q)+V_{qq}(q,Q)} \qquad \text{IV.A.8}$$

From the SOCs for the maximization of $\pi(q,Q)$ it follows that the curve $q=q^{\square}(Q)$ implied by $\lambda(q,Q)=0$ is steeper than the curve $q=q^{\bullet}(Q)$ implied by $\pi_q(q,Q)=0$ and that the sign of the slopes of both curves equals the sign of $F_q(q)+V_q(q,Q)-QV_{qQ}(q,Q)$; see Fig. IV.A.4 and Fig. IV.A.5.

$F_q(q)+V_q(q,Q)<QV_{qQ}(q,Q)$ $\qquad$ $F_q(q)+V_q(q,Q)>QV_{qQ}(q,Q)$

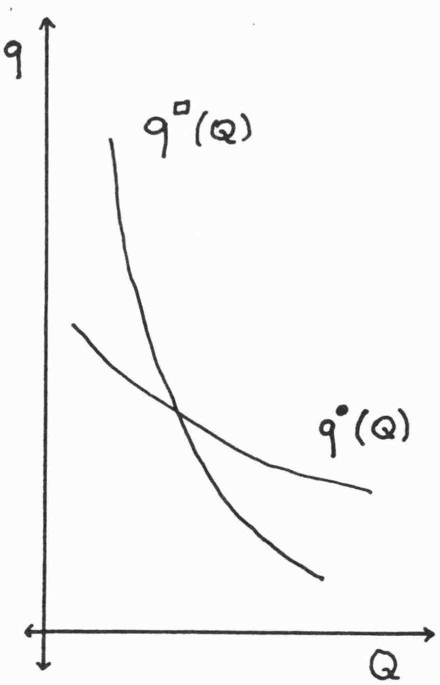


Fig. IV.A.4

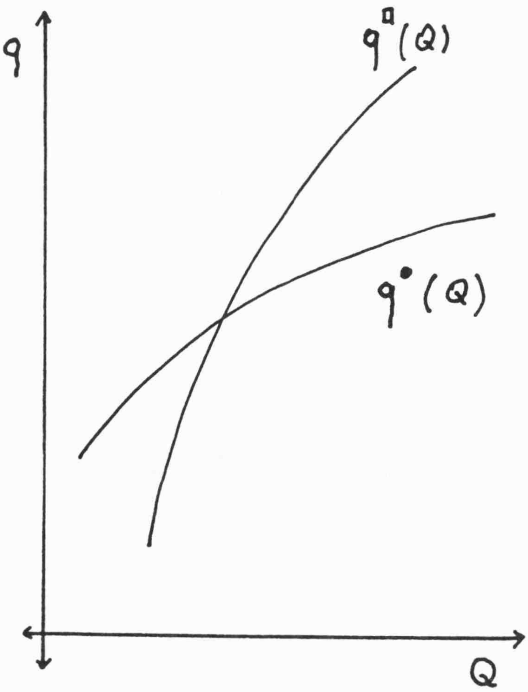
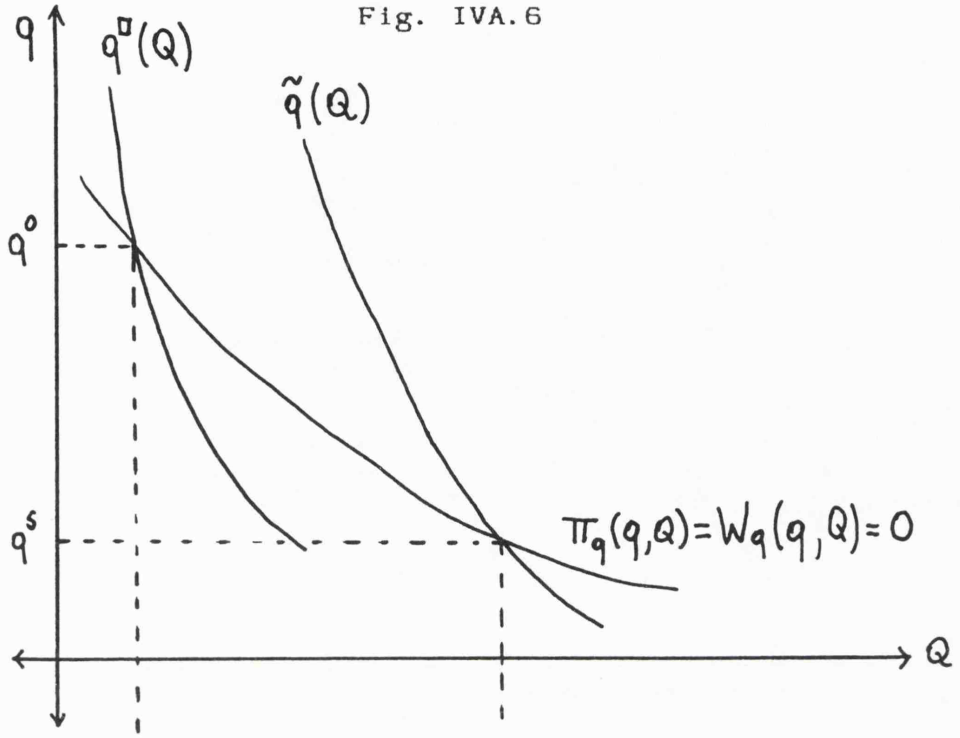


Fig. IV.A.5

To compare the welfare-maximizing solution (q^o, Q^o, n^o) with the free-entry oligopoly equilibrium (q^s, Q^s, n^s), notice that from IVA.1 and IVA.6 it follows that $q=q^o(Q)$ lies everywhere below the curve $q=\tilde{q}(Q)$ (assuming that $F_q(q)+V_q(q,Q)<QV_{qQ}(q,Q)$; the opposite case is symmetrical) as shown in Fig. IVA.6.

Similarly, it is easy to establish that for $q^o \geqslant q^s$, $\pi_Q(q^o, Q, n)=0$ lies everywhere below $W_Q(q^s, Q, n)=0$ and that for $q^o>q^s$, $\pi(q^o, Q, n)=0$ lies everywhere above $W_n(q^s, Q, n)=0$. These results are summarized in Fig. IVA.7 which establishes the claim made in the text (IV.21-23) that $n^o>n^s$, $Q^o<Q^s$ and that $q^o \gtrless q^s$ according as to whether $F_q(q)+V_q(q,Q) \lessgtr QV_{qQ}(q,Q)$.

Fig. IVA.6



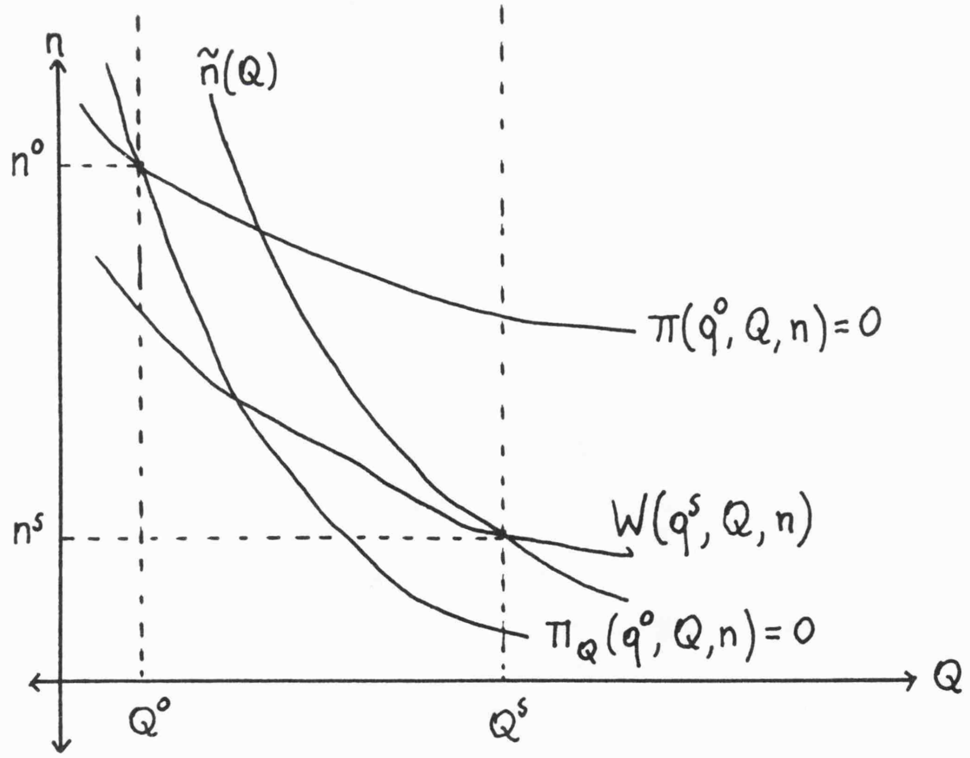


Fig. IVA.7

REFERENCES

Arrow, K., (1962), "Economic Welfare and the Allocation of Resources for Inventions", in R.R. Nelson (ed.) ***The Rate and Direction of Inventive Activity***, Princeton, Princeton University Press, 1962.

Baldwin, W. and Scott, J., (1987), ***Market Structure and Technological Change***, Chur: Harwood Academic Publ, 1987.

Brander, J. and Spencer, B., (1983), "Strategic Commitment with R&D: the Symmetric Case", ***Bell Journal of Economics***, vol. 14, 1983, pp. 225-35.

Brander, J. and Spencer, B., (1985), "Tacit Collusion, free entry and welfare", ***Journal of Industrial Economics***, vol. 33 (3), pp. 277-294.

Commission of the European Communities (1988), "Proposal for A Council Directive on the Legal Protection of Biotechnological Inventions", Brussels, 17 Oct. 1988.

Beath, J., Katsoulacos, Y., and Ulph, U., (1989), "The Game-Theoretic Analysis of Innovation: A Survey", ***Bulletin of Economic Research***, 1989, pp.

Dasgupta, P. and Heal, G., (1979) ***Economic Theory and Exhaustible Resources***, Cambridge, Cambridge University Press, 1979.

Dasgupta, P., and Stiglitz, J., (1980), "Industrial Structure and the Nature of Innovative Activity", ***Economic Journal***, 90, pp. 266-293.

Dasgupta, P. and Stiglitz J., (1980), "Uncertainty, Industrial Structure and the Speed of R&D", ***Bell Journal of Economics***, 11 (1), 1980, pp. 1-28.

Friedman, J., (1983), ***Oligopoly Theory***, Cambridge, Cambridge University Press, 1983.

Fudenberg, D., and Tirole, J., (1986), "Noncooperative Game Theory for Industrial Organization : An Introduction and Overview", Univ. of California, Dept. of Economics, Working Paper No. 8613.

Gilbert, R., and Shapiro C.,(1988), "Optimal Patent Breadth and Length", Working Paper, University of California, Berkeley, October 1988.

Green, J., and Scotchmer, S., (1989), "Technological Licensing and the Novelty Requirement in Patent Law", Working Paper, Hoover Institution, Stanford University, Aug. 1989.

Grossman, G., and Shapiro, C., (1987), "Dynamic R&D competition", ***Economic Journal***, 97, 1987, pp. 372-87.

Horsley, A., (1982), "A New Vector Space Theory of Microeconomics", mimeo.

Horstmann, I., MacDonald, G., and Slivinski, A., (1985), "Patents as Information Transfer Mechanisms: To Patent or (Maybe) Not to Patent, *Journal of Political Economy*, 1985, 93(5), pp. 837-58.

Jaskold Gabszewicz, J. and Thisse, J. (1979), "Price competition, quality, and income disparities", *Hournal of Economic Theory*, 20, 1979, pp. 340-359.

Jaskold Gabszewicz, J. and Thisse, J. (1980), "Entry (and exit) in a differentiated industry", *Journal of Economic Theory*, 22, 1980, pp. 327-338.

Kamien, M., and Schwartz, N., (1982), *Market Structure and Innovation*, Cambridge: Cambridge University Press,, 1982.

Kitch, E., (1977), "The Nature and Function of the Patent System", *Journal of Law and Economics*, 20, 1977, pp. 265-90.

Klemperer, P., (1989), "How Broad Should the Scope of Patent Protection Be?", Working Paper, Oxford University Institute of Economics and Statistics, Feb. 1989.

Klette, T. and de Meza, D., (1986a) "Is the Market Biased Against Risky R&D?", *Rand Journal of Economics*, 17, pp. 133-139.

Klette, T. and de Meza, D., (1986b) "Major Inventions Require the Least Patent Protection", Mimeo.

La Manna, M., (1986), "Why Are Profits Correlated with Concentration", *Economics Letters*, 22, 1986, pp. 81-85.

La Manna, M., (1987), "A Note on Quality, Market Structure, and the Importance of the 'Integer Constraint'", *Bulletin of Economic Research*, 39, 1987, pp. 303-307.

La Manna, M., Macleod, R., and de Meza, D., (1989), "The case for permissive patents", *European Economic Review*, vol. 33 No. 7, Sept. 1989, pp. 1427-43.

La Manna, M., (1990), "Patents as Early Warning Signals: the effects of alternative patent regimes on industry structure and social welfare", in *Market Structure and Strategy*, Proceedings of the International Conference on Market Structure and Strategy, Dundee, Aug. 1988, edited by George Norman and Alec Gee, London: Macmillan, forthcoming.

Lee, T., and Wilde, L., (1980) "Market Structure and Innovation: A Re-formulation", *Quarterly Journal of Economics*, 94, pp. 429-436.

Levin, R., (1986), "A New Look at the Patent System", *American Economic Review* (Papers and Proceedings) 76, pp. 199-202.

Loury, G,, (1979), "Market Structure and Innovation", *Quarterly Journal of Economics*, 93, pp. 395-410.

Nelson, R.R., (1982), "The Role of Knowledge in R&D Efficiency", *Quarterly Journal of Economics*, 97, 1982, pp. 453-70.

Machlup, F., and Penrose, E., (1950), "The Patent Controversy in the Nineteenth Century", ***Journal of Economic History,*** 10(1), 1950, pp. 1-29.

Nordhaus, W., (1969), ***Invention, Growth, and Welfare***, Cambridge, Mass.: MIT Press, 1969.

Plant, A., (1934), "The Economic Theory Concerning Patents for Invention", ***Economica,*** 1, pp. 31-51.

Reinganum, J. (1985), "A Two-Stage Model of Research and Development With Endogenous Second-Mover Advantages", ***International Journal of Industrial Organization***, 1985, 3, pp. 275-92.

Salant, S., Switzer, S., and Reynolds, R., (1983), "Losses from Horizontal Merger: the effects of an exogenous change in industry structure on Cournot-Nash Equilibrium", ***Quarterly Journal of Economics,*** 98, 1983, pp. 185-199.

Scherer, F.M., (1972),"Nordhaus' Theory of Optimal Patent Life: A Geometrical Reinterpretation", ***American Economic Review,*** 1972, pp. 422-427.

Schmalense, R., (1979), "Market Structure, Durability, and Quality: A Selective Survey", ***Economic Inquiry,*** vol. 17, pp. 177-196.

Seade, J. (1980), "On the effects of entry", ***Econometrica,,*** vol. 48 (2), pp. 479-489.

Shaked, A. and Sutton, J., (1982), "Relaxing price competition through product differentiation", *Review of Economic Theory,* 49, 1982, pp. 3-13.

Shaked, A. and Sutton, J., (1983), "Natural Oligopolies", *Econometrica,* 41, 1983, pp. 1469-1484.

Sheshinski, E., (1976), "Price, Quality and Quantity Regulation in Monopoly Situations", *Economica,* vol. 43, pp. 127-137.

Stoneman, P., (1987), *The Economic Analysis of Technology Policy,* Oxford, Clarendon Press, 1987.

Tandon, P., (1982), "Optimal Patents with Compulsory Licensing", *Journal of Political Economy,* vol. 90, 1982, pp. 470-86.