

First mover advantages and optimal patent protection

F. M. Scherer

Published online: 25 February 2015
© Springer Science+Business Media New York 2015

Abstract This paper examines the economic logic underlying so-called “first mover advantages” as alternatives to patent protection in allowing innovators protection from immediate imitation and time to recoup their research and development investments. These include a head start, lower costs due to learning by doing, and product-differentiating reputational advantages. William Nordhaus’ theory of optimal patent protection is extended to cases of product, as distinguished from cost-saving, innovation. Over the parameters covered by a simulation analysis, profit potentials sufficient to induce R&D investment are found much more frequently than R&D incentive failures. R&D failures occur most commonly in small markets and situations in which imitators’ erosion of the innovator’s market share is particularly rapid. Implications for patent policy are drawn.

Keywords Technological innovation · Intellectual property or patents · Product differentiation

JEL Classification O34

Patents or similar exclusive privileges have been awarded for many centuries to encourage invention and innovation. For early history, see Machlup (1958) and Kaufer (1989). Absent some barrier to competitive imitation such as patent rights, the underlying theory holds, competition might materialize so rapidly that the inventor-innovator is unable to recoup the investment made in effecting its innovation. In a pioneering theoretical contribution, Nordhaus (1969) derived conditions showing the social welfare-maximizing life of patent grants. Virtually ignored in both the Nordhaus theory of patent protection and rationales underlying patent laws has been another thrust of the literature. Empirical studies have shown repeatedly that on average, but with notable exceptions, patent protection is a relatively unimportant requisite for business firms’ investment in research, development,

F. M. Scherer (✉)

John F. Kennedy School of Government, Harvard University, Cambridge, MA 02138, USA
e-mail: mike_scherer@hks.harvard.edu

and innovation. Much more important in the average case are diverse non-patent advantages from being the first to commercialize a new product or process. Eliciting estimates from 25 British companies, Taylor and Silberston (1973) found that having all their patents subject to compulsory licensing at “reasonable” royalties would on average reduce the firms’ research and development expenditures by only 8 %. Surveying 100 US companies, Mansfield (1986) reported that the weighted average number of inventions actually introduced by respondents that would not have been developed had *no* patent protection been available was roughly 14 %.¹ See also Scherer et al. (1959, Chapter 12), Levin et al. (1987), Cohen et al. (2004), and Graham et al. (2009). This paper seeks to advance the theory of patent protection by quantifying approximations to the “first mover advantages” that sustain investment in invention and innovation without formal patent protection.

1 The elementary theory

The elementary logic underlying the grant of temporary but exclusive patent rights on *product inventions* is illustrated by Fig. 1. The demand for a potential invention is D_1 . The marginal cost of production (excluding front-end R&D costs) is assumed constant at $C-MC$. Marketing its product under conditions of monopolistic competition, the innovator equalizes marginal revenue (not shown) with marginal cost and sets price OA , earning a net profit before deduction of sunk R&D costs (more accurately called a quasi-rent) of AC per unit and a total quasi-rent given by rectangle $ABEC$. If this continues for a sufficiently long period, i.e., with patent protection, the discounted present value of the quasi-rents will cover and (the innovator hopes) exceed the original R&D investment, and the investment will prove to be profitable. But if others can readily imitate the innovator’s product, its demand will shift to D_2 in period 2 and D_3 in period 3 (arrows), etc., with quasi-rents shrinking to $abeC$ in period 2 and $a'b'e'C$ in period 3, etc. With such rapid entry, the discounted present value of quasi-rents could be less than the original R&D investment. If the would-be innovator foresees this, no R&D investment will be forthcoming. Again, patents inhibit imitative entry and hence encourage investment in research, development, and innovation.

If however the innovator enjoys non-patent first mover advantages,² competitive imitation may be delayed even in the total absence of patent rights. These are of several forms.³

For one, it takes time for would-be imitators to recognize the advantages of an innovation and quite possibly even more time to carry out their own technological work needed to imitate successfully. In some cases, when the imitator can benefit from knowledge spillovers, that expense may be much less than the first mover’s expense, but in other cases (such as developing new airliners) the imitator may have to spend as much and take as much time as the first mover did.⁴

Second, the innovator may be able to keep important details of its underlying technology secret, inhibiting imitation. This is more likely for process (i.e., internal cost-saving)

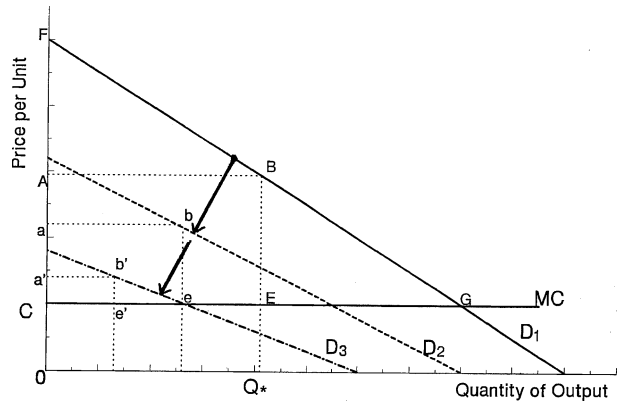
¹ The foregone invention counts are weighted by aggregate research and development expenditures in the surveyed industry groups.

² The term probably originated in game theory.

³ This summary follows Scherer (1980, pp. 444–447).

⁴ Levin et al. (1987, p. 809) found that the R&D cost needed to duplicate a major *unpatented* new product exceeded 50 % of the first mover’s R&D in 86 percent of the surveyed industries.

Fig. 1 How limitation can reduce the innovator's profit



innovations—those analyzed by Nordhaus—than for product imitations, but even for new products, non-obvious production tricks may have to be discovered and mastered.

Third, and very importantly, the first to market a new product often engrains in the minds of consumers an “image” of superiority—that is, a product differentiation advantage—allowing it to retain a substantial market share while charging prices appreciably higher than those realizable by latecomers.⁵

Fourth, in industries such as aircraft, semiconductors, and solar converters, unit production costs fall with additional production and hence “learning by doing.” The first mover begins progressing down its learning curve sooner than others and may therefore enjoy a substantial cost advantage over latecomers.⁶

Finally, economies of scale in production or marketing may require that a market be tightly oligopolistic, with only a few sellers contending for position, among other things through product innovation and differentiation. High R&D costs required for innovation may reinforce this structural condition. And with well-established marketing channels, the first mover can expect to retain preferential access to customers accustomed to patronizing particular sales representatives and/or retailers unless it falls significantly behind the product quality of rivals. In this case, a firm may be confident that when it innovates, it can retain at least a substantial share of the market after rivals imitate. However, it must also fear that if actual or potential rivals are the first innovative movers, they will capture, perhaps permanently, its own market share. In this case, characterizable as Schumpeterian “creative destruction” (Schumpeter 1942, Chapter 7), companies are impelled to invest in innovation by the threat of competition, whether or not patent protection can be anticipated.⁷

⁵ For early recognition of the reputation effect, see Plant (1934, p. 44). Seminal empirical evidence on this theme came from research with Strategic Planning Institute data. See Buzzell and Gale (1987), Buzzell and Farris (1977), and Robinson and Fornell (1985). See also Bond and Lean (1977) and Schmalensee (1982). On how first movers with an image or cost advantage can, through strategic pricing, impede the rate at which rivals enter their market, see Gaskins (1970), especially pp. 42–53; and more generally, Scherer (1980, pp. 232–243).

⁶ On how Intel used this advantage strategically, see Scherer (2011, pp. 50–54).

⁷ The theory of such R&D races can be traced back to Scherer (1967), which deliberately neglected the role of patent protection. That dominant firms often but not always respond with a lag to rival technological challenges is shown by the seven case histories presented in Scherer (2008), Chapter 44. See also Baldwin and Scott (1987), and, (introducing a new term, “disruptive” innovation) Christensen (1997).

2 The economic theory of optimal patent duration

The leading theory on how the duration of patent grants affects investment in research and development and how patent lives can in turn be adapted to maximize a broader conception of social welfare was originated by Nordhaus (1969).⁸ Nordhaus focuses on what is best called *process innovation*, that is, advances in technology reducing the cost of production (the MC line in Fig. 1) and thereby increasing the size of the innovator's profit (rectangle ABEC in Fig. 1). He calls inventions that merely reduce marginal cost without inducing an output expansion *run of the mill inventions*; those that also increase output are called *drastic inventions*. Nordhaus devotes no attention to possible pre-patent issue competition among multiple firms to achieve cost reductions through invention. His assumption, to which counter-arguments will be recognized later, is analogous to what Kitch (1977) has called the *prospect theory of innovation*.

Nordhaus argues that the amount of cost reduction achieved is systematically and positively related through an *invention possibility function* to the amount of research and development conducted: the more R&D, the higher the percentage cost reduction. Given its invention possibility function, and given a payoff time structure determined by the government's patent life policy choice, that is, the period over which the inventor can exploit its invention without competition, the inventing firm is assumed to choose the amount of cost-reducing R&D that maximizes the discounted present value of its invention-dependent profits. Nordhaus shows that the longer the patent's life is, the more cost-reducing R&D will be induced, all else equal. In determining how long patents should have their exclusionary power, government policy-makers in turn are assumed by Nordhaus to choose a patent life that maximizes the invention's contribution to *social welfare*, including not only the profits achieved through cost reductions by the inventing firm but also increases in consumers' surplus realized during the life of the patent (only in drastic invention cases) plus those like triangle BGE in Fig. 1 realized when patent protection ends and competition forces prices down to the new level of marginal cost.

Nordhaus' pioneering contribution clarified relationships that had previously been visualized at best qualitatively and imprecisely. Compare e.g., Machlup (1958, pp. 66–73). It had, however, five significant limitations. First, it focused on cost-saving or process innovations which, statistics available at the time showed, amounted to only about a fourth of all research and development expenditures incurred by American industries. The remaining three-fourths comprised R&D directed toward creating or improving *products* that would eventually be sold to consumers or other companies. To be sure, what for the inventing firm is a product (e.g., a computer-controlled lathe or a turbojet engine) may for the purchasing firm be a cost-reducing process, but the market dynamics, we shall see, are different.⁹ Second, it assumed implicitly that patents were the only barrier to competitive imitation of inventions, ignoring the first mover advantages that were at the time beginning to be recognized as barriers to rapid imitation. Third, as is customary in mathematical economics, it focused on achieving optimal first-order conditions, with marginal private or social benefits being equalized to corresponding marginal costs. In so doing, it deemphasized the possibility that many alternative outcomes might be profitable for inventing firms and improve social welfare, even though they do not achieve a maximum

⁸ A much earlier, more qualitative but prescient contribution was by Sir Arnold Plant (1933). A geometric interpretation of the Nordhaus theory is offered in Scherer (1972).

⁹ On the flow of technologically advanced products to become processes in purchasing industries, see Scherer (1982).

maximorum. Fourth, by assuming a single inventor investing in new technology without recognizable technological rivals, it in effect accepted the prospect theory of invention and rejected the alternative and empirically plausible rent-seeking theory, with profoundly different implications for patent policy.¹⁰ And fifth, it ignored the fact that invention is a cumulative process, and in particular, that how patent rights affect the use of an invention made at time t can significantly affect the further progress of technology at time $t + n$.¹¹

This paper seeks to fill lacunae left by the first three of these Nordhaus assumptions. The fourth and fifth, requiring much richer and more diverse empirical foundations, are commended to others.

3 Theoretical foundations

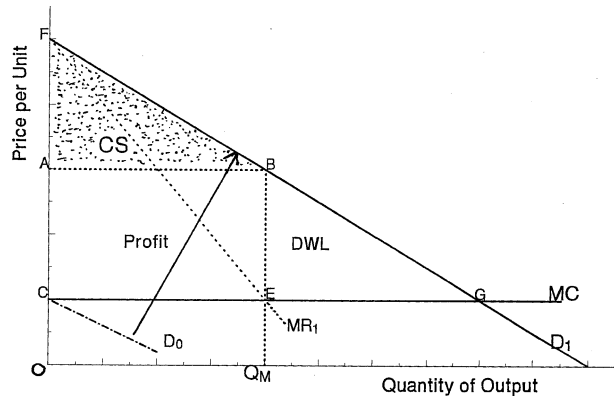
Our focus here is on product innovations. The most prominent attribute of product innovations is that they make products more attractive to consumers, shifting demand functions upward and outward, perhaps making it profitable to produce and market a product that, without the invention, would not have been commercially viable. Demand functions reveal the quantity of a given product that will be demanded at diverse alternative prices, holding constant, as is conventional in partial equilibrium analyses, the second-order adjustment of substitute or complementary products' prices. Given this, we emphasize here the consequences for economic welfare within the immediate innovation-impacted product market, ignoring spillovers in possibly related product markets.

We assume that successful research and development shifts the relevant product's demand function upward so that it lies above marginal cost (whose level may also be affected by the invention project) and hence that profitable commercialization is possible. The model is illustrated at its simplest in Fig. 2. We assume that marginal cost is constant per unit produced at level C-MC. Without invention, the relevant demand function is D_0 , which lies at all points but the zero quantity value below marginal cost and hence leaves no possibility of profitable production. Invention shifts the demand function upward by CF (arrow), with the new demand function D_1 allowing profitable production. Assuming that the innovator has monopoly power in pricing its product, albeit taking into account the prices of potential substitute products, the innovator derives its marginal revenue function MR_1 , chooses to market an output of OQ_M at price OA, and realizes a profit (strictly defined, quasi-rent) above its variable costs given by the rectangle ABEC. This profit is counted as a benefit from a broader society-wide perspective. But in addition, and in contrast to the Nordhaus run-of-the-mill case, sales of the new product yield to buyers a consumers' surplus measured by the dot-shaded triangle CS. The emergence of competition can have either or both of two effects. For one, imitators capture some of the innovator's sales, squeezing its demand function to the left and perhaps (depending upon demand elasticities and the dynamics of rivalry) forcing the innovator to reset its price, leading to reduced profits. The assumption accepted here is that the innovator continues to have some monopoly power (i.e., under conditions of differentiated product oligopoly), so that it still faces a downward-sloping (though changed) marginal revenue function and can at least for a while after entry set a price above its marginal cost. Imitative rivals too are assumed to be differentiated product oligopolists tacitly cooperating, *a la* Chamberlin, at least within limits, to set prices that maximize joint oligopoly profits. Second, the

¹⁰ See Barzel (1968) and McFetridge and Rafiquzzaman (1986).

¹¹ See e.g., Merges and Nelson (1990) and Scotchmer (1991).

Fig. 2 How innovation shifts product demand function outward



alternative assumption, implicitly accepted in the Nordhaus model, is that when patents or first mover advantages are lost, unrestrained price competition breaks out immediately and prices are driven all the way down to marginal cost MC . In this case, what was previously potential but lost consumers' surplus measured by triangle BGE (labelled DWL) in Fig. 2 is transformed into actual consumers' surplus—a measurable social gain. We will take this result as a benchmark in our analysis of socially optimal outcomes. At the same time, firms' quasi-rents are transformed into consumers' surplus on what is conventionally assumed to be a 1:1 basis.¹²

In our model, the vigor of the innovator's product R&D effort determines (without stochastic variation, important in the real world) how far the innovator's demand function is thrust upward. What is needed then is an analogue to Nordhaus' invention possibility function. We focus on the resultant height of the demand function, leaving for our simulation analysis of alternative cases the demand function's slope and hence the breadth of the market. An intuitive rationale for this assumption is that product innovation affects consumers' willingness to pay, which is measured by a demand function's vertical dimension. The horizontal dimension is more closely related to the scope of the relevant market, which might arguably be said to be more exogenous than endogenous. Specifically, we assume that the demand function is shifted through product innovation from an intercept at point C in Fig. 2, leaving sales intrinsically unprofitable, to an endogenous intercept point F , where F can vary from identity with C (if no R&D is performed) to five times OC ($S = 5$). Initial experiments assumed an exponential shift function $S = RD^k$, with $k < 1$ to imply diminishing marginal returns. This approximates the approach taken by Nordhaus with his invention possibility function. Those experiments, however, yielded implausibly high shift values and hence market sizes. Therefore, a quadratic approximation

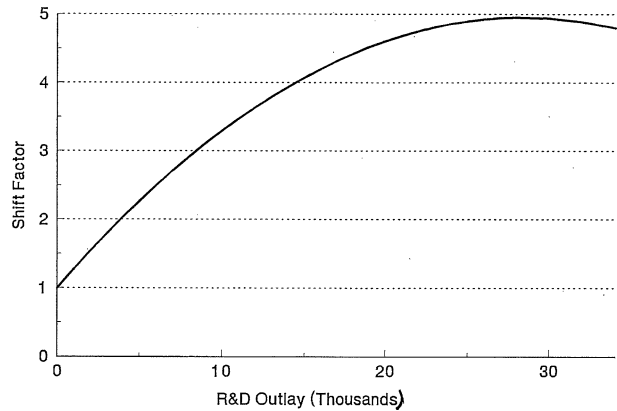
$$S = 1 + .2784 RD - .0049(RD)^2 \quad (1)$$

was used, with R&D outlays measured in two digits only, e.g., in thousands of dollars (or millions, for larger projects).¹³ It is illustrated in Fig. 3. Its maximum value is realized at R&D expenditures of roughly 28 (000), with $S = 4.95$ at that value.

¹² For an alternative view, see Scherer (2004).

¹³ In Scherer (1972), I suggested more inflected cubic or in one case linear dogleg invention possibility functions.

Fig. 3 Relationship of demand shift factor to R&D expenditure



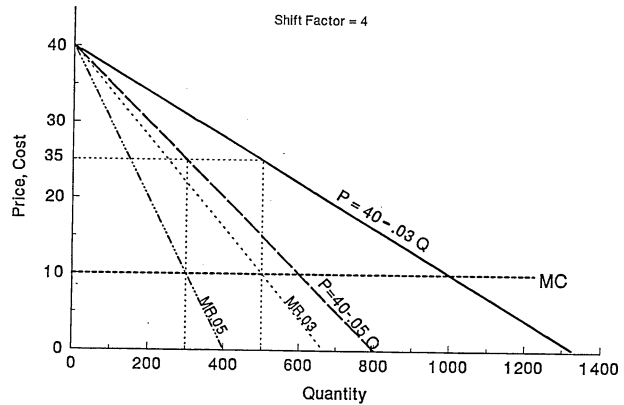
Needless to say, even with a given demand intercept shift, markets can be of widely varying sizes. This arguably exogenous variability was taken into account by assuming demand functions to be linear in price, with equation $P = 10S - (\text{slope}) Q$ and slopes dP/dQ varying from $-.03$ to $-.07$. Two such cases are illustrated in Fig. 4. With linear demand and a given vertical intercept, the demand curves are iso-elastic, i.e., with the same price elasticity of demand for any given vertical level. This allows an assumption, simplifying calculation in computer simulations, that profit-maximizing prices are identical (\$35 per unit in Fig. 4) for linear demand functions with the same vertical intercept but varying slopes. The numerical assumptions to be sure determine the outcomes, and exploration with alternative parameter sets is encouraged. But the assumptions permit a wide range of circumstances to be investigated.

4 The effect of differing patent lives revisited

For broad insight into how outcomes vary with alternative market parameters and patent lives, payoff matrices were calculated for three benchmark cases: with demand slopes of $-.03$, $-.05$, and $-.07$. In a preliminary stage, shift variables were determined as a function of R&D outlays varying by units from 1 to 35. For each R&D outlay and R&D-determined shift variable, profit-maximizing prices, quantities, and (given constant marginal cost of \$10 per unit) price–cost margins were computed. The product of price–cost margins times profit-maximizing quantities yielded the annual quasi-rent following from any given R&D expenditure. The quasi rents are assumed for simplicity to be constant over time (an assumption varied later). For each possible patent life (by twos) from 4 years to infinity, the discounted present value of quasi-rents realized by the innovator was calculated. Patent lives were assumed to begin at year 0. Quasi-rents were assumed to begin flowing in at year 1, i.e., after a year's R&D and production setup. For each patent life, a discount factor associated with the years of sale under patent protection was applied to the annual quasi-rent estimate. The assumed discount rate was 12 % per annum.¹⁴ Following conventional assumptions (but inconsistent with the theory and evidence on first mover advantages to be

¹⁴ Nordhaus (1969, p. 81) assumes a discount rate of 20 %. A lower rate was chosen here mainly because an expert panel retained by the US Office of Technology Assessment (1993, pp. 267–283) found 12 % to be an empirically-supported risk-adjusted real discount rate for research-intensive pharmaceutical firms.

Fig. 4 Graph of demand and cost functions



addressed later), the calculations assumed that when patents expire, quasi-rents fall precipitously to zero and remain there subsequently. From the discounted present values so calculated, front-end research and development costs (multiplied from column (1) by 1000) were subtracted to yield the discounted present value from a year 0 vantage point of profits net of R&D outlays—i.e., an indication of whether any given R&D expenditure was in the net present value-enhancing.

The private payoff matrices computed under these assumptions are too unwieldy to report in full. Here we provide in Table 1 only one example, for the intermediate $-.05$ demand slope case. Other tables are available at <http://ssrn.com/abstract=2538621>. The first column (labelled “Rand”) contains alternative R&D expenditures (in thousands of dollars). The second column reports the annual quasi-rent (labelled “Profit”) resulting from each R&D expenditure, given the shift variable’s impact on the demand curve’s intercept value and the resulting profit-maximizing solution, assuming marginal cost to be \$10 per unit. The quasi-rent is assumed constant per year from the time of product introduction to the time of patent expiration, after which it plunges to zero. The remaining columns report discounted present profit values (with front-end R&D costs deducted) for various patent lives. In the third column (headed “dpvt04”, for discounted present value through year 4, implying a 4-year patent life), for example, we see that discounted present values are negative for all R&D expenditures. Positive values begin only at patent lives of 6 years (column 4) or higher. The maxima among these positive values are printed in bold face. One sees that the profit-maximizing R&D expenditures increase with longer patent lives until they stabilize beyond a 12-year life.¹⁵

Given the $-.05$ demand slope assumed in Table 1, the internal rate of return from the profit-maximizing R&D investment (Rand = 22) and a patent life of 16 years is 27%.¹⁶ With a larger market (slope $-.03$), the IRR is approximately 49%; with the smallest of the three markets analyzed (slope $-.07$), the IRR is 21.5%. Obviously, for less lucrative cases, the internal rate of return is at least 12% over all payoff matrix cells with positive net present values.

¹⁵ For the largest of the three assumed markets (slope $-.03$), the optima stabilize beyond patent lives of 12 years and for the smallest (slope $-.07$) beyond 18-year lives.

¹⁶ The calculations were done assuming finite annual discount terms (e.g., $(1/(1+r)^t)$) rather than the e^{-rt} integrals used in computing the column values for various patent lives.

Table 1 Private payoff matrix for diverse patent lives; medium-size market (demand slope = $-.05$)

Rand	Profit	dpvt04	dpvt06	dpvt08	dpvt10	dpvt12	dpvt14	dpvt16	dpvt18	dpvt20	infin
1	37	-916	-875	-843	-817	-797	-782	-769	-760	-752	-724
2	144	-1678	-1519	-1394	-1296	-1218	-1158	-1110	-1072	-1043	-934
3	313	-2301	-1957	-1686	-1473	-1305	-1173	-1070	-988	-924	-687
4	536	-2803	-2213	-1749	-1385	-1098	-872	-694	-555	-445	-40
5	806	-3199	-2313	-1615	-1067	-635	-296	-29	181	347	956
6	1116	-3506	-2278	-1313	-553	45	515	885	1176	1405	2248
7	1460	-3738	-2132	-868	125	907	1522	2006	2387	2686	3789
8	1831	-3909	-1894	-310	937	1917	2689	3296	3773	4148	5532
9	2223	-4032	-1586	338	1852	3043	3979	4716	5296	5752	7432
10	2631	-4121	-1226	1052	2843	4252	5361	6233	6919	7458	9447
11	3049	-4187	-832	1807	3883	5516	6801	7812	8606	9232	11,536
12	3472	-4242	-421	2584	4948	6807	8270	9421	10,326	11,038	13,662
13	3895	-4296	-11	3360	6012	8098	9739	11,030	12,046	12,844	15,788
14	4314	-4361	384	4118	7055	9365	11,182	12,612	13,736	14,621	17,881
15	4723	-4446	750	4838	8054	10,584	12,574	14,139	15,370	16,338	19,908
16	5120	-4559	1074	5505	8991	11,733	13,890	15,587	16,922	17,971	21,841
17	5500	-4710	1342	6102	9847	12,793	15,110	16,933	18,367	19,494	23,651
18	5861	-4905	1543	6615	10,605	13,744	16,213	18,155	19,683	20,885	25,314
19	6198	-5151	1667	7031	11,251	14,570	17,181	19,235	20,851	22,122	26,806
20	6509	-5456	1705	7338	11,770	15,256	17,998	20,155	21,852	23,186	28,105
21	6791	-5824	1647	7525	12,149	15,787	18,648	20,899	22,669	24,061	29,194
22	7043	-6262	1487	7583	12,378	16,150	19,118	21,452	23,288	24,732	30,055
23	7262	-6773	1217	7503	12,447	16,337	19,396	21,803	23,696	25,185	30,674
24	7447	-7360	833	7278	12,347	16,336	19,473	21,941	23,882	25,409	31,037
25	7595	-8028	328	6902	12,072	16,140	19,340	21,857	23,837	25,395	31,135
26	7707	-8779	-300	6370	11,617	15,744	18,991	21,545	23,554	25,134	30,959

Table 1 continued

Rand	Profit	dpvt04	dpvt06	dpvt08	dpvt10	dpvt12	dpvt14	dpvt16	dpvt18	dpvt20	infin
27	7780	-9615	-1055	5679	10,976	15,143	18,421	20,999	23,028	24,622	30,503
28	7815	-10,536	-1938	4827	10,147	14,333	17,626	20,216	22,254	23,856	29,763
29	7812	-11,544	-2950	3812	9130	13,314	16,605	19,194	21,231	22,832	28,736
30	7770	-12,639	-4090	2634	7924	12,085	15,359	17,933	19,959	21,552	27,424
31	7689	-13,819	-5359	1296	6530	10,649	13,888	16,436	18,441	20,017	25,828
32	7571	-15,083	-6754	-201	4953	9008	12,197	14,706	16,680	18,232	23,954
33	7415	-16,430	-8272	-1853	3195	7167	10,291	12,748	14,682	16,202	21,806
34	7225	-17,857	-9908	-3655	1263	5133	8176	10,570	12,454	13,935	19,395
35	6999	-19,360	-11,659	-5601	-836	2913	5862	8181	10,006	11,441	16,731

Two main insights emerge at this benchmarking stage. First, over a wide range of parameter values and patent lives, investment in research and development is profitable, even if not present value-maximizing. Second, R&D markets may fail totally—i.e., no investment in R&D will be made—if short patent lives are coupled with relatively small market spaces.

5 Constrained social welfare maximization *a la* Nordhaus

What has been presented thus far is an analysis of how profitable R&D investments are from the perspective of private enterprises for various market parameters and patent lives. Now we ask, how do these results mesh with the interests of the broader society in which firms operate? In other words, what constellation of parameters and choices maximizes *social* welfare? Social welfare here is measured by the discounted sum of surpluses from innovation realized by both the enterprises that provide the innovations *and* the consumers who utilize them. To explore this, a new kind of payoff matrix is needed. Our previous analysis focused on *private* innovating firm payoffs only—that is, the discounted present value of quasi-rent rectangles like ABEC in Fig. 2 minus the (undiscounted) value of R&D expenditures. Social payoffs are measured during the patent's life by the rectangle ABEC *plus* the triangle of consumers' surplus FBA in Fig. 2. But after patents expire, values do not plunge to zero as assumed in the previous analysis. Rather, the profit rectangle ABEC is converted into consumers' surplus as price competition drives prices down to marginal cost; triangle FBA in Fig. 2 continues to be realized as surplus by consumers (in industrial product cases, users) of the product; and in addition, triangle BGE (labelled DWL) now becomes consumers' surplus as lower prices induce increased consumption by previously under-served consumers. Given the linearity assumptions underlying our demand and cost analysis, the pre-patent-expiration social surplus is exactly 1.5 times the innovator's quasi-rent, while the post-expiration surplus is 2.0 times the original innovator's quasi-rent. Each such surplus must, of course, be discounted to present value in the final reckoning, here by assumption at a constant 12 % discount rate.

The resulting payoff matrices are not reproduced here (but see <http://ssrn.com/abstract=2538621>). One finds that except for very low R&D outlays, social surpluses are positive across all patent lives between 4 and 20 years (because additional consumer surpluses are recognized). Since discounted social surpluses vary only modestly with either short or long patent lives, welfare-maximizing levels of R&D for each assumed patent life entail almost uniform R&D investments. To learn, following Nordhaus, which *patent life* maximizes discounted *social* value, given profit-maximizing behavior by private innovators, one identifies in the relevant social payoff matrix the *private* payoff-maximizing R&D outlays for diverse patent lives, choosing the one associated with a patent life for given market parameters that maximizes discounted *social* returns. Following this procedure, one finds that social welfare-maximizing patent lives increase systematically from six to eight to 14 years, the smaller the relevant market (of the three alternatives analyzed) is. This is an important insight. One sees also that the choices maximizing discounted private profits involve systematically lower R&D expenditures than those that maximize social welfare. This is the natural result of the fact that many of the benefits driving social welfare maximization are external to private company decision-makers.

How large is the difference between socially optimal and private profit-maximizing R&D outlays over varying market parameters? With the medium-size market of Table 1 (slope = $-.05$), private R&D expenditures at the socially optimal patent life of 8 years are

22(000), or about 84.6 % of the socially optimal R&D expenditure 26(000) (not ascertainable without inspection of the underlying social payoff matrix). Repeating this analysis for the alternative demand parameters assumed, we obtain the following percentages:

	Demand slope	Private/social R&D percent
Smallest market	-0.07	88.0
Intermediate market	-0.05	84.6
Largest market	-0.03	88.9

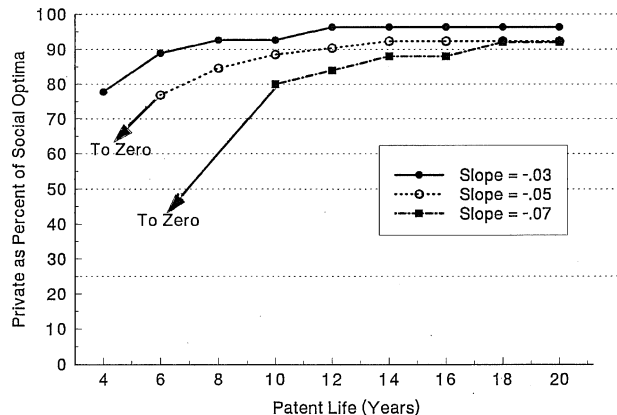
These are for the patent lives and hence R&D expenditure levels that maximize social welfare for the three assumed market sizes. A broader analysis is provided by Fig. 5, which arrays private/social divergences over a range of patent lives from 4 to 20 years. With the largest market, i.e., demand slope = -0.03, the divergence is smallest. With smaller markets, the divergence rises, plummeting to R&D of zero—that is, a complete R&D market failure—for patent lives less than 6 years in the medium-sized market case and patent lives less than 10 years in the smallest market. The implication is that the patent system makes its strongest incremental contribution to sustaining innovation in relatively small markets—those in which, it should be noted, the social gains from innovation are smallest.

6 Alternative first mover assumptions

We advance now to territory that remains unexplored. This paper takes off from the recognition that barriers other than patents impede rapid imitation of a first mover’s commercialized invention. Not all of the first mover advantages articulated earlier can be treated analytically. We proceed by incorporating the following simplified assumptions.

To model recognition lags, secrecy, and the imitator’s need to perform its own R&D, we test the effect of simple time lags ranging from 2 to 4 years following the innovator’s product innovation date at year 1—that is, with imitation in years 3, 4, and 5.

Fig. 5 Private R&D optima as percent of social optima for varying patent lives (in years)



A second key assumption is that the innovator's loss of market share to imitators is not instantaneous, but because of the first mover's image and cost advantages, imitators' market penetration rates are constrained. We assume concretely that the innovator's market share decays exponentially at alternative rates of 10, 15, and 20 % per year.

That rivals do not gain market position instantaneously with new products suggests a symmetric assumption, contrary to the naive assumptions underlying our previous Nordhaus-like analysis, that first movers must also build up their patronage at an incremental rate rather than capturing the whole market in the first year after a new product's debut. Specifically, we assume for simplicity here that the first mover penetrates the relevant product market at the rate of 50 % per annum, or more precisely, that the share of the market it does not capture, ignoring imitation, decays at the rate of e^{-pt} , where p is the 0.5 penetration coefficient and t is a running year variable with value of zero at the time of innovation.¹⁷ Needless to say, this assumption disfavoring early year quasi-rents reduces the discounted present values of innovator profits to values much less than those assumed in the previous patent life analyses. Not surprisingly, trial simulations of first mover effects without this innovator penetration assumption (not reported here) yielded appreciably higher discounted innovator profits than those presented here. Analyses with the 0.5 penetration rate are emphasized here because of their greater believed realism.

Figure 6 summarizes the innovator market share implications of the alternative penetration and imitator erosion scenarios, assuming that imitation begins in year 3, the quickest of the three imitation lag scenarios. With no imitation, the innovator starting at the beginning of year 2 achieves 95 % of its market potential after 6 years, i.e., in year 7. But erosion by imitators causes an increasing loss of innovator market position and leads within a decade to substantially atrophied innovator shares—as low as 3–8 % within 20 years.

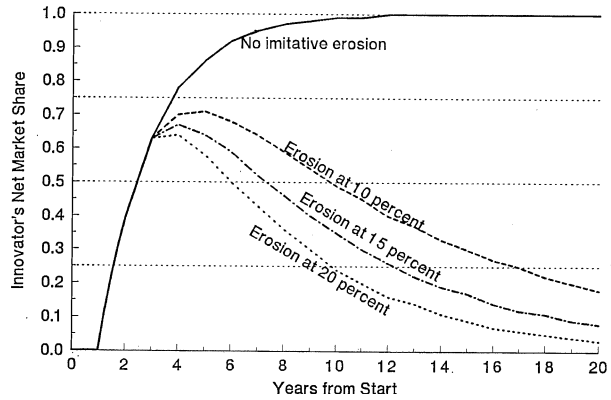
That patent-free imitation might within the life span of conventional patents (assumed absent in our first mover model) drive innovator market shares to very low values clashes with the notion that the first mover gains significant “image” advantages, permitting it to maintain premium prices indefinitely and to combine that product differentiation advantage with pricing strategies that impede competitive entry. This consideration underlies an additional set of simulations in which the first mover's retained market share is constrained not to fall below 30 %. The 30 % assumption was drawn from Bond and Lean (1977) and pioneering research done in the 1970s using rich data collected by the Strategic Planning Institute.

Table 2 illustrates the results of simulations embodying our first mover assumptions (but no 30 % market share constraint) with a market of medium size, i.e., with slope of $-.05$, and diverse initial imitation lags and imitator penetration rates. As with Table 1, the first column contains alternative R&D outlays “Rand” (in thousands) and the second “Profit” the resulting (constant) annual quasi-rent potential tapped by the innovator *and* its imitators.¹⁸ The remaining column headings disclose the first year of imitation (e.g., in column (3), with imitation after 2 years, i.e., in the third year analyzed); and “ero...” reports the rate at which imitators erode the innovator's profit, e.g., in column (3), at a

¹⁷ The combination of penetration and imitator erosion coefficients mirrors the approach taken in Scherer (1967), equation [7].

¹⁸ It is assumed that imitators' products do not enlarge or reduce the size of the total quasi-rent pool; e.g., because the number of imitators is small enough to retain a joint profit-maximizing oligopoly. Note a key difference between Tables 1 and 2: the innovator quasi-rents in Table 2 from which discounted present values are calculated encompass not the total market potential, but only that part of the potential tapped and/or retained by the innovator.

Fig. 6 Innovator's market share under alternative limitation scenarios; innovator captures market at 50 % annual rate



10 % annual rate. The numbers in those columns are discounted present values of innovator quasi-rents less front-end R&D costs. For each such column, the maximum discounted present value entry, if positive, is printed in bold face.

With the assumed medium market size of Table 2, all nine scenarios yield at least one positive discounted present quasi-rent value, although with the earliest imitator entry in year 3 and a high erosion rate of 20 %, there is only one tiny positive entry. More generally, the range of unprofitable investment alternatives (minus signs) is smaller, the longer the imitation lag and the lower the imitator erosion rate. With the smallest of the assumed markets, slope = $-.07$, catastrophic R&D failure appears. *No* investment level is profitable with imitation beginning 2 years after innovation, i.e., in year 3. With a 3-year imitation lag, positive profits occur only with the lowest erosion rate of .10. A 4-year imitation lag (i.e., under the heading “Imitation in Year 5”) is more successful, with positive profit outcomes for some .10 and .15 erosion rates (implying privately optimal R&D levels of 19(000) or 20(000), even if not for the more rapid erosion at 20 %. Clearly, market size, imitation lags, and erosion rates are critical to the success of first mover advantages in providing effective incentives alternative to the patent system.

Constraining the innovator's market share loss to imitators so that a minimum share of 30 % is retained, consistent with the early pharmaceutical rivalry histories studied by Bond and Lean (1977), improves the incentive picture. All scenarios yield some positive discounted profit-maximizing outcome for the largest market. In the smallest market, R&D failure occurs again with higher erosion rates and imitation lags of 2 or 3 years, but with a 4-year imitation lag, positive equilibria are found for all three erosion rates. The effect of first mover advantages in allowing innovators to defend at least a minority share of their new markets through economies of scale, cost advantages, and the possibility of sustaining profitable price differentials contributes significantly to incentives for R&D investment.

Tables 3 and 4 summarize the impact of diverse first mover scenarios on the strength of R&D investment incentives. As a benchmark, we take the level of R&D investment that would maximize social welfare with 16-year patent lives, drawn from analyses like that of Table 1. Table 3 reports on 27 simulations, assuming no floor to the innovator's market share following imitation. Across the nine outcomes for varying imitation lags and erosion rates, we find for the largest market that profit-maximizing R&D outlays with imitation average 90.1 % of the social welfare-maximizing levels. For the medium-size market, they average 81.6 % of the social optima. Averages conceal more than they reveal in the smallest of the three assumed markets. Six out of nine cases are failures, with no profitable

Table 2 Payoff matrix with penetration and imitation lags; medium-size market (demand slope = $-.05$)

Rand	Profit	Imitation in year 3				Imitation in year 4				Imitation in year 5					
		im3ero10		im3ero15		im4ero10		im4ero15		im5ero10		im5ero15		im5ero20	
1	37	-842	-868	-885	-830	-853	-868	-818	-839	-853					
2	144	-1391	-1490	-1558	-1342	-1432	-1492	-1298	-1378	-1432					
3	313	-1679	-1894	-2041	-1574	-1767	-1899	-1478	-1651	-1769					
4	536	-1738	-2107	-2357	-1558	-1889	-2115	-1394	-1690	-1892					
5	806	-1598	-2152	-2530	-1327	-1825	-2165	-1080	-1525	-1830					
6	1116	-1289	-2056	-2579	-914	-1603	-2074	-571	-1188	-1610					
7	1460	-838	-1841	-2524	-347	-1249	-1865	101	-705	-1258					
8	1831	-271	-1530	-2387	344	-786	-1559	906	-105	-798					
9	2223	385	-1143	-2184	1133	-241	-1179	1815	587	-255					
10	2631	1107	-702	-1933	1992	366	-744	2799	1346	349					
11	3049	1872	-225	-1652	2897	1013	-273	3832	2148	993					
12	3472	2657	269	-1355	3824	1680	214	4889	2972	1657					
13	3895	3442	764	-1058	4752	2346	702	5947	3795	2321					
14	4314	4209	1243	-775	5659	2995	1174	6982	4600	2967					
15	4723	4938	1690	-520	6526	3608	1615	7975	5366	3578					
16	5120	5613	2093	-303	7334	4172	2011	8905	6077	4138					
17	5500	6218	2436	-137	8067	4670	2349	9754	6717	4634					
18	5861	6739	2709	-33	8709	5089	2616	10,507	7270	5051					
19	6198	7162	2901	1	9246	5418	2802	11,147	7724	5377					
20	6509	7476	3000	-45	9664	5643	2897	11,660	8065	5601					
21	6791	7669	2999	-179	9952	5757	2891	12,035	8284	5713					
22	7043	7732	2889	-407	10,100	5749	2777	12,260	8370	5703					
23	7262	7656	2663	-735	10,098	5612	2547	12,325	8314	5565					
24	7447	7435	2314	-1170	9938	5339	2196	12,222	8109	5290					

Table 2 continued

Rand	Profit	Imitation in year 3			Imitation in year 4			Imitation in year 5		
		im3ero10	im3ero15	im3ero20	im4ero10	im4ero15	im4ero20	im5ero10	im5ero15	im5ero20
25	7595	7062	1839	-1714	9615	4924	1719	11,945	7750	4874
26	7707	6532	1233	-2373	9123	4363	1111	11,487	7231	4313
27	7780	5843	493	-3147	8459	3653	370	10,845	6548	3602
28	7815	4991	-382	-4039	7619	2791	-507	10,016	5700	2741
29	7812	3976	-1395	-5050	6603	1777	-1519	8999	4684	1727
30	7770	2798	-2544	-6180	5410	611	-2668	7793	3502	560
31	7689	1458	-3829	-7427	4043	-707	-3951	6401	2155	-757
32	7571	-42	-5247	-8790	2504	-2173	-5368	4825	644	-2222
33	7415	-1697	-6796	-10,266	796	-3784	-6914	3070	-1025	-3833
34	7225	-3503	-8471	-11,851	-1074	-5537	-8585	1142	-2848	-5584
35	6999	-5453	-10,266	-13,541	-3100	-7424	-10,377	-953	-4819	-7469

Table 3 Analysis of first mover R&D incentives

Erosion rate	Imitator begins entry at year		
	3	4	5
Case 1: large market (demand slope = $-.03$)			
.10	92.6 %	92.6 %	92.6 %
.15	88.9	88.9	92.6
.20	85.2	88.9	88.9
Case 2: medium-size market (slope = $-.05$)			
.10	84.6 %	86.5 %	88.5 %
.15	78.9	80.8	84.6
.20	73.1	76.9	80.8
Case 3: small market (slope = $-.07$)			
.10	0	80.0 %	80.0 %
.15	0	0	76.0
.20	0	0	0

First mover's DPV-maximizing R&D outlays as percentage of social benefit-maximizing outlay. Social optimum computed with 16-year patent life. (First mover enters market at 50 % rate.)

Table 4 Analysis of first mover incentives, with 30 % innovator market share floor

Erosion rate	Imitator begins entry at year		
	3	4	5
Case 1: large market (demand slope = $-.03$)			
.10	92.6 %	92.6 %	92.6 %
.15	88.9	92.6	92.6
.20	88.9	90.7	92.6
Case 2: medium-size market (slope = $-.05$)			
.10	84.6 %	88.5 %	88.5 %
.15	80.8	84.6	88.5
.20	80.8	84.6	84.6
Case 3: small market (slope = $-.07$)			
.10	76.0	80.0 %	84.0 %
.15	0	0	80.0
.20	0	0	76.0

First mover's DPV-maximizing R&D outlays as percentage of social benefit-maximizing outlay. Social optimum computed with 16-year patent life. (First mover enters market at 50 % rate.)

level of R&D expenditure. For the remaining three cases, the average is 78.7 %. Turning to Table 4, we see that a floor under the innovator's eventual market share improves outcomes slightly. For the largest market, the capped versus social average is 91.6 %; for the medium-size market, 85.1 %. For the smallest of the three markets, there is also an improvement: five of the nine cases yield positive levels of R&D investment.

These averages reveal that first mover advantages fall short of yielding socially optimal levels of R&D investment. But the patent system is also an imperfect allocator. From our initial benchmark analysis, we found that Nordhaus-optimal R&D investment levels are 88.9 % of the social optimum in the largest of the three markets, 84.6 % in the medium-size market, and 88.0 % in the smallest market.¹⁹ Over those three, the simple average is

¹⁹ Within the range of Nordhaus-optimal values, socially optimal R&D levels remain essentially constant (subject to rounding error attributable to considering only integer values) across a wide range of patent lives, including our 16-year benchmark.

87.2 %. Ignoring the cases of R&D failure, the average across the three markets with first mover advantages substituting for patent protection and letting innovator market shares fall at most to 30 % is 86.3 %—only 1 % below the benchmark with-patent optima averages. We conclude at least tentatively, subject to confirmation with richer simulations, that incentive systems focused on first mover advantages do a quite tolerable job of allocating resources to research and development. Their principal weakness, at least by the assumptions underlying our analysis, is proneness to zero-R&D corner solutions, especially in small markets.

The incidence of innovation failure also suggests an interesting though less general insight. Out of the 54 simulations summarized by Tables 3 and 4, 10, or 18.5 %, entail failures to invest at all in R&D because expected quasi-rents fall short of R&D costs over all alternative levels of R&D investment. This is not greatly at odds with Mansfield's (1986) finding, from hypothetical questions posed to 100 US corporations, that in the total absence of patent protection, 14 % of the innovations they actually commercialized would not have been made. In other words, first mover advantages alone would have provided insufficient incentive.

7 Discussion

The analysis presented here is at best a too long-delayed first step. Much more remains to be done. First mover advantages could be modelled in different and richer ways. And the basic model itself is highly simplified, focusing on a single clearly-designated inventor obtaining (or not obtaining) one patent, not hundreds or thousands, protecting an invention targeted toward a specific node in product characteristics space.²⁰ A more complex and more realistic analysis could deal with innovators and presumed later movers offering differentiated products, each occupying its own niche in product characteristics space.²¹ The underlying mathematics would be more complex but in principle tractable, at least for a subset of economists excluding the author. Multi-year R&D and nonlinear R&D—payback structures should be explored.²² Also needed is additional empirical work illuminating the various types of first mover advantages, their impact on the retention of market share and profits, and the rates at which both innovators and imitators penetrate their target markets. Levin et al. (1987), Cohen et al. (2004), and Graham et al. (2009) provide a start, but much more could be learned.

A more important simplification here has been the assumption, following Nordhaus, that there is a single identifiable first mover whose invention, absent patent protection, is then imitated by one or more other firms. This scenario corresponds most closely to what Edward Kitch (1977) has called “the prospect theory of invention.” In an alternative conception, changes in the forces of supply (i.e., the state of scientific and technological knowledge) and demand endogenously spur numerous firms to invest rivalrously in R&D in the hope of winning a preferred market position, perhaps protected by patents. The

²⁰ If multiple patents are required to sustain full monopoly, our early Nordhaus-like assumption that a single patent suffices is wrong. If so, the more the additional patents bring the innovator to a full monopoly, the closer to the Nordhaus case reality comes.

²¹ See Scherer (1980, pp. 393–398) for the author's own survey of a literature pioneered by others. A significant methodological contribution is Hausman et al. (1994).

²² See Ravenscraft and Scherer (1982). That analysis, like the insights in this paper on reputation effects, was developed using Strategic Planning Institute data.

rivals' efforts may in the limiting case be so vigorous that the sum of R&D costs incurred by all participants equals or even exceeds the realizable pool of quasi-rents, so that in hindsight R&D investment approximates a zero-profit equilibrium. For a pioneering statement of this "rent-seeking" model, see Barzel (1968).²³ Re-analysis by McFetridge and Rafiqzaman (1986, Table 1) revealed that when the implicit prospect assumption is replaced by a rent-seeking assumption, with R&D costs escalating to exhaust quasi-rents, the welfare-maximizing patent life over 32 sets of parameter values averaged 21.2 years following the strict Nordhaus assumptions and 0.8 years when rents are totally dissipated by competing inventors. Somewhere between these unrealistic extremes is the messier world of oligopoly, where investment in research and development is less a quest for protected niches in profitable product space than a Schumpeterian struggle for survival against disruptive market inroads by product-differentiating rivals.

Perhaps more serious a simplification is our neglect of uncertainty in assuming that the R&D investments supported are those with positive discounted present values. There is abundant research showing that R&D payoff projections are not only uncertain, but that the distribution of profit outcomes is quite skew, with a long thin tail containing the projects that ultimately yield the lion's share of returns.²⁴ Skewness in turn makes it difficult to use classical hedge strategies, that is, by supporting a portfolio of R&D projects. In the first mover simulations performed here, with only three alternative demand parameters, the innovator's DPV maxima ranged from zero (10 cases) to a maximum of 38,200, with a mean of 11,900, a median of 8,000, and a standard deviation of 12,400. This is skew, but not nearly as skew as the distributions studied by the Scherer and Harhoff (2000). It is plausible that larger, well-established companies approach their R&D decisions with something at least approximating a portfolio mentality, expecting the less frequent good results to compensate or indeed over-compensate for the disappointments. This is less likely for smaller and especially startup ventures, but the behavioral consequences, though empirically fragile, could be surprising. Startups might be more risk-averse on average, requiring especially strong patent protection and/or first mover advantages to justify R&D investment decisions. But there is also reason to suppose that at least in the United States, small technological innovators are on average risk-lovers, taking long-shot investments that, if they could be quantified in advance, yield returns actuarially less than "normal." See Scherer (2001). What is certain is that both theories and policies on how R&D investments are motivated must take uncertainty into account.

7.1 Policy implications

Given this standard economist's caveat that further research is needed, it might be premature to suggest policy implications. But the questions addressed here are important and almost totally ignored in past policy deliberations. Therefore, it is useful to sketch some directions in which policies might be adapted.

We begin from the widely accepted premise that all is not well in the realm of patent law and its administration.²⁵ Patents have proliferated without any evident increase in the

²³ For applications to pharmaceutical R&D revealing dramatic differences in implications, depending upon whether prospect or rent-seeking model assumptions are followed, see Scherer (2004, 2010, pp. 562–569).

²⁴ The pioneering empirical analysis was Mansfield et al. (1977). Evidence on skewness is presented *inter alia* in Scherer and Harhoff (2000).

²⁵ See UK Commission (2002), US National Research Council (2004), Jaffe and Lerner (2004), and Bessen and Maurer (2008).

rate of technological innovation; and patent litigation—sometimes escalating to patent wars—has imposed high costs upon participants, governmental adjudicative bodies, and technical progress. These problems suggest that fundamental policy changes should be seriously considered.

In view of evidence that non-patent first mover advantages often provide sufficient incentive for technological innovation even without patent protection, the logical step would be scaling back significantly the normal scope and duration of patent protection while retaining the possibility of long-lived patents for identifiable special cases. One starting point would be to limit routine grants for the protection of invention to short lives, like the German Gebrauchsmuster (petty patents), with streamlined granting procedures and effective lives limited to 5 years.²⁶ A 5-year patent would both complement alternative first mover advantages and provide a brief lag in its own right—a lag found important to improved profitability in the first mover cases analyzed here as imitation lags were increased from 2 to 4 years. The law could then identify two kinds of exceptions: a general class attempting to encompass the cases in which first mover advantages are systematically inadequate, and a special administrative procedure to grant longer-term protection when patent applicants provide persuasive evidence that conventional incentives are insufficient.

The most serious R&D failures, our analysis suggests, occur when markets are too small to support substantial R&D investments. Given the uncertainties that pervade R&D investment, identifying small-market exceptions administratively in advance is undoubtedly not feasible. A workable surrogate would be to limit full-term (i.e., 20 year) patent grants to unaffiliated independent inventors and companies with, say, fewer than 20 employees at the time of patent application.²⁷ Newness and smallness are presumably correlated with modest market prospects. A more fine-tuned but administratively feasible policy (using data like those employed in competition agencies' merger analyses) would withhold full-term patents mainly for firms with substantial (i.e., >10 %) market shares in Census-defined industries that are structured oligopolistically, i.e., in which the four leading sellers originate more than 40 % of national market sales or value added.²⁸ For such firms, established channels of distribution and brand images provide non-patent first mover advantages, and the threat of Schumpeterian creative destruction normally generates R&D incentives at least as potent as those offered by the prospect of a strong patent position contingent upon success.

Exceptions to these general presumptions could then be handled through expedited case-by-case administrative procedures. For example, applicants might seek extended protection by showing that (1) the costs of R&D are extraordinarily high compared to benchmarks in the relevant line of business;²⁹ (2) the underlying R&D is susceptible to quick replication at costs much lower than those incurred by the original inventor;³⁰ (3) a full-term patent issuing from the relevant application will be licensed at "reasonable" royalties and without other restrictive provisions to all good-faith license seekers; and (4)

²⁶ Needless to say, given the 20-year patent mandate incorporated in the 1994 TRIPS Treaty of Marrakesh, the change would have to be world-wide and not confined to a single nation. But if the negotiators for TRIPS got the balance wrong in the first place, their choices should indeed be reversed. Compare Ryan (1998).

²⁷ It would be necessary to restrain "sham" transactions, e.g., when a large firm finances the R&D of small entities and then acquires the resulting patents with their full-term rights.

²⁸ Such data have been published at roughly 5-year intervals by the US Census Bureau.

²⁹ For guidance, see Levin et al. (1987, p. 809).

³⁰ This is commonly the case in pharmaceuticals. See (Scherer 2010, pp. 560–561).

(often accompanying (3)), that the patent-seeking firm conducts R&D only and makes its profits by licensing the resulting technology to other manufacturers or users.³¹ To be sure, administrative costs at the Patent Office would be increased by the need to adjudicate such petitions for exception. But if the default case is reduced to a Gebrauchsmuster-type patent, the Patent Office's costs of search to show that minimal standards of inventiveness are met and that patent issue is not barred by the existence of prior art—functions that are both costly and error-prone—could be reduced substantially.

These proposals are not intended to be anything like the last word on a long-controversial subject. Rather, they are articulated to stimulate debate on an important point that has been virtually ignored in past analyses—the undeniable fact that non-patent first mover advantages serve an imitation-delaying function like that exercised by patents, permitting investors in invention and innovation to anticipate monetary rewards for their efforts that, even without patents, are often sufficient to make the investments worth while.

References

- Baldwin, W. L., & Scott, J. T. (1987). *Market structure and technological change*. Chur: Harwood.
- Barzel, Y. (1968). Optimal timing of innovations. *Review of Economics and Statistics*, 50, 348–355.
- Bessen, J., & Maurer, M. (2008). *Patent failure*. Princeton: Princeton University Press.
- Bond, R., & Lean, D. (1977). *Sales, promotion, and product differentiation in two prescription drug markets*. Washington: Federal Trade Commission Staff Report.
- Buzzell, R. D., & Farris, P. (1977). Marketing costs in consumer goods industries. In H. Thorelli (Ed.), *Strategy + structure = performance* (pp. 122–145). Bloomington: Indiana University Press.
- Buzzell, R. D., & Gale, B. (1987). *The PIMS principles: Linking strategy to performance*. New York: Free Press.
- Christensen, C. (1997). *The innovator's dilemma*. Boston: Harvard Business School Press.
- Cohen, W. M., et al. (2004). Protecting their intellectual assets: Appropriability conditions and why U.S. manufacturing firms patent (or not). Cambridge, MA: National Bureau of Economic Research Working Paper no. 7552.
- Gaskins, D. W. (1970). *Optimal pricing by dominant firms*, Ph.D. Dissertation, University of Michigan.
- Graham, S. J. H., et al. (2009). High technology entrepreneurs and the patent system: Results of the 2008 Berkeley patent survey. *Berkeley Technology Law Journal*, 4, 1255–1327.
- Hausman, J., et al. (1994). Competitive analysis with differentiated [sic] products. *Annales d'Economie et de Statistique*, 34, 159–180.
- Jaffe, A., & Lerner, J. (2004). *Innovation and its discontents*. Princeton: Princeton University Press.
- Kaufers, E. (1989). *The economics of the patent system*. Chur: Harwood.
- Kitch, E. (1977). The nature and function of the patent system. *Journal of Law and Economics*, 20, 265–290.
- Levin, R. C., et al. (1987). Appropriating the returns from industrial research and development. *Brookings Papers on Economic Activity*, 3, 783–832.
- Machlup, F. (1958). *An economic review of the patent system*, Study No. 15 of the U.S. Senate Subcommittee on Patents, Trademarks, and Copyrights.
- Mansfield, E. T. (1986). Patents and innovation: An empirical study. *Management Science*, 32, 173–181.
- Mansfield, E. T., et al. (1977). *The production and application of new industrial technology*. New York: Norton.
- McFetridge, D. G., & Rafiqzaman, M. (1986). The scope and duration of patent rights and the nature of research rivalry. *Research in Law and Economics*, 8, 91–129.
- Merges, R., & Nelson, R. R. (1990). On the complex economics of patent scope. *Columbia Law Review*, 90, 839–916.

³¹ Classic examples in famous US patent rights law include radio circuit designer Hazeltine Research (*Automatic Radio v. Hazeltine Research*, 339 U.S. 827 (1950)); and glass bottle machinery designer Hartford-Empire (*U.S. v. Hartford-Empire Co. et al.*, 323 U.S. 386 (1947)). See also the description of semiconductor designer ARM's policies in Moore (2013), especially pp. 56–57. Licensing practices would presumably continue to be subject to competition policy rules.

- Moore, J. F. (2013). *Shared purpose*. Privately published.
- Nordhaus, W. D. (1969). *Innovation, growth, and welfare*. Cambridge, MA: MIT Press.
- Plant, A. (1934). The economic theory concerning patents for inventions. *Economica*, 1, 30–51.
- Ravenscraft, D., & Scherer, F. M. (1982). The lag structure of returns to research and development. *Applied Economics*, 14, 603–620.
- Robinson, W. T., & Fornell, C. (1985). Sources of marketing pioneer advantages in consumer goods industries. *Journal of Marketing Research*, 22, 305–317.
- Ryan, M. (1998). *Knowledge diplomacy: Global competition and the politics of intellectual property*. Washington: Brookings.
- Scherer, F. M. (1967). Research and development resource allocation under rivalry. *Quarterly Journal of Economics*, 81, 359–394.
- Scherer, F. M. (1972). Nordhaus' theory of optimal patent life: A geometric reinterpretation. *American Economic Review*, 62, 422–427.
- Scherer, F. M. (1980). *Industrial market structure and economic performance* (2nd ed.). Chicago: Rand-McNally.
- Scherer, F. M. (1982). Interindustry technology flows in the United States. *Research Policy*, 11, 227–245.
- Scherer, F. M. (2001). The innovation lottery. In R. Dreyfuss, et al (Eds.), *Expanding the boundaries of intellectual property* (pp. 3–22). Oxford: Oxford University Press.
- Scherer, F. M. (2004). A note on global welfare in pharmaceutical patenting. *The World Economy*, 27, 1127–1142.
- Scherer, F. M. (2008). Technological innovation and monopolization. In W. D. Collins, et al (Eds.), *Issues in competition law and policy* (vol. II, pp. 1033–1068). Chicago: American Bar Association.
- Scherer, F. M. (2010). Pharmaceutical Innovation. In B. Hall & N. Rosenberg (Eds.), *Handbook: Economics of innovation* (pp. 539–574). Amsterdam: North Holland.
- Scherer, F. M. (2011). Abuse of dominance by high technology enterprises: A comparison of U.S. and E.C. approaches. *Journal of Industrial and Business Economics*, 38, 39–62.
- Scherer, F. M., & Harhoff, D. (2000). Technology policy for a world of skew-distributed outcomes. *Research Policy*, 29, 559–566.
- Scherer, F. M., Herzstein, S., et al. (1959). *Patents and the corporation*, 2nd ed. Boston: Privately published.
- Schmalensee, R. (1982). Product differentiation advantages of pioneering brands. *American Economic Review*, 72, 349–365.
- Schumpeter, J. A. (1942). *Capitalism, socialism, and democracy*. New York: Harper.
- Scotchmer, S. (1991). Standing on the shoulders of giants: Cumulative research and the patent law. *Journal of Economic Perspectives*, 5, 29–41.
- Taylor, C. T., & Silberston, Z. A. (1973). *The economic impact of the patent system*. Cambridge: Cambridge University Press.
- U. S. Office of Technology Assessment. (1993). *Pharmaceutical R&D: Costs, Risks, and Rewards*. Washington.
- United Kingdom Commission on Intellectual Property Rights. (2002). *Integrating Intellectual Property Rights and Development Policy*, London.
- U.S. National Research Council. (2004). *A Patent System for the 21st Century*, Washington.