Appropriability of innovation benefit as a predictor of the source of innovation *

ERIC VON HIPPEL

Alfred P. Sloan School of Management, MIT, Cambridge, MA 02139, USA

Final version received October 1981

It has been empirically observed that, in some industries product users are the most frequent sources of product innovations while, in other industries, product manufacturers are. I hypothesize that such differences are caused by differences in the ability of these two "functional" categories of innovators to appropriate innovation benefit. I explore this hypothesis by examining the real-world effectiveness of mechanisms (such as patents and lead time) used for the appropriation of innovation benefit and the dependence of this effectiveness on the functional relationship between innovator and innovation.

1. Introduction

Empirical studies of the functional locus of innovation, the variable modeled in this paper and first studied by Peck [1] categorize innovators in terms of the *functional* relationship via which they derive benefit from the innovations they create. Thus, if one is studying a sample of process machinery innovators, those who use the innovative machinery in production would be grouped in terms of that functional relationship into a "user" category, innovators who benefit economically from manufacturing the process machinery innovations grouped into a "manufacturer" category, etc.

The functional locus of innovation has proven very useful in innovation research because it is reliably measurable and because it often displays very strong differences between samples examined. Thus, we see from table 1 that Berger [5] and Boyden [6] find that 100% of their samples of, respectively, engineering polymer innovations and polymer additive innovations were developed by manufacturers of these. In sharp contrast, Lionetta [7] and von Hippel [9] find users to be the developers of 85% and 68% respectively of the samples of process machinery innovations whose antecedents they investigated.

The striking differences empirically observed in the functional locus of innovation are doubtless a function of several variables. In this paper, however, I explore the hypothesis that such differences can be effectively modeled in terms of one variable only: the different abilities of would-be innovators holding different functional relationships to a given innovation to appropriate benefit from that innovation. [1] More specifically, I hypothesize that the functional locus of innovation can be effectively modeled in terms of appropriability of innovation benefit if and as three conditions hold in the real world, namely, would-be innovators: (1) *are not* able to capture benefit from non-embodied knowledge characterizing their innovations: (2) *are* able to capture benefit from output-embodied knowledge relating to their innovations; and (3) differ significantly in their *ability* to capture benefit from output-embodied innovation knowledge. In the following sections of this paper I identify and explore the real-world effectiveness of mechanisms available to innovators for the appropriation of innovation benefit, and provide an initial empirical test of the proposed model.

* The research reported on in this paper was supported by the National Science Foundation under Grant no. PRA 77-07830.

[1] Readers interested in a more general discussion of appropriability of innovation benefit may wish to refer to key papers by Arrow [10], Nelson [11], and Pakes and Schankerman [12], in addition to the pioneering paper by Peck [1].

Research Policy 11 (1982) 95–115
North-Holland Publishing Company

0048-7333/82/0000-0000/$02.75

Table 1
Empirical data on the functional source of commercialized industrial innovations

Study	Nature of innovations and sample selection criteria	n	Innovation developed [a] by User (%)	Mfr (%)	Other (%)
Knight [2]	Computer innovations 1944-1962:				
	- system reaching new performance high	143	25	75	
	- systems with radical structural innovations (level 1)	18	33	67	
Enos [3]	Major petroleum processing innovations	7	43	14	43 [b]
Freeman [4]	Chemical processes and process equipment available for license, 1967	810	70	30	
Berger [5]	All engineering polymers developed in US after 1955 with >10 mm pounds produced in 1975	6	0	100	
Boyden [6]	Chemical additives for plastics: all plasticizers and UV stabilizers developed post World War II for use with four major polymers	16	0	100	
Lionetta [7]	All pultrusion processing machinery innovations first introduced commercially 1940-1976 which offered users a major increment in functional utility	9	89	11	
von Hippel [8]	Scientific instrument innovations:				
	- first of type (e.g. first NMR)	4	100	0	
	- major functional improvements	44	82	18	
	- minor functional improvements	63	70	30	
von Hippel [9]	Semiconductor and electronic subassembly manufacturing equipment.				
	- first of type used in commercial production	7	100	0	
	- major functional improvements	22	63	21	16 [c]
	- minor functional improvements	20	59	29	12 [c]

[a] Attribution of an innovation to a user or manufacturer "developer" is determined by which of these first builds and utilizes the innovation in conformance with his economic function. Thus, attribution to a user source is made if a user builds and *uses* an innovation before a manufacturer builds and sells a commercial version. And conversely, attribution to a manufacturer source is made if a manufacturer builds and *sells* a commercial version of an innovation before a user builds and uses a home-made version. NA data excluded from percentage.
[b] Attributed to independent inventors/invention development companies.
[c] Attributed to joint user-manufacturer innovation projects.

2. The ability to predict the functional locus of innovation as a function of the appropriability of innovation-related benefit

The economic benefits which an innovator might obtain from his innovation can be segregated into two mutually exclusive and jointly exhaustive categories: (1) benefit from "output-embodied" knowledge and (2) benefit from non-embodied knowledge. Benefit from output-embodied knowledge is obtained by an innovator via in-house use of his innovation in his product and/or process and the consequent embodiment of its value in the output of his firm. Benefit from non-embodied knowledge is obtained by an innovator from the sale or licensing of non-embodied knowledge regarding his innovation to others. Let us consider whether we would logically expect to be able to predict the functional locus of innovation – i.e. the functional relationship of innovator to innovation – under each of two extreme cases regarding the ability of an innovator to capture benefit from his innovation:

Case 1 Total ability to capture benefit from output-embodied knowledge and total ability to capture benefit from non-embodied knowledge.

Case 2 Total ability to capture benefit from output-embodied knowledge but *no* or only an imperfect ability to capture benefit from non-embodied knowledge.

2.1. Predictions regarding the functional locus of innovation under case 1 conditions

If we assume that an innovator has "perfect", costlessly enforceable property rights to his innovation, i.e. if, without cost to himself, he can totally control its diffusion and capture benefit from innovation users, manufacturers, and others to the point where adoption becomes a matter of indifference to them, then the benefits capturable by an innovator would be the same no matter what his own functional relationship to the innovation at issue. Thus, under case 1 conditions we can make no prediction regarding the functional locus of innovation on the basis of appropriability of benefit considerations.

The reasoning behind the above conclusion is that costless [2] enforcement of property rights would allow any innovator to set the fees charged to each innovation beneficiary, and each class of beneficiaries, so as to attain the maximum return. The role which the innovator himself happens to play with regard to the innovation – user, manufacturer, etc. – does not influence his fee-setting decision because he is equally able to capture innovation returns from his own company and other companies. This being so, he has no incentive to concentrate benefits in his own company even if the direct return from the particular innovation can be "leveraged" [3] by its user to create larger "other returns" over time.

2.2. Predictions regarding the functional locus of innovation under case 2 conditions

Under case 2 conditions we assume that: (1) the innovator has temporary monopoly power over the innovation information embodied in his output and thus is able to capture significant benefit from embodying that knowledge in the output of his firm; and (2) the innovator has *no* or only a very imperfect ability to capture benefit from diffusing non-embodied information regarding his innovation to others. Faced with this situation, the economically rational firm, seeking to maximize its joint return from output-embodied knowledge and non-embodied knowledge, would wish to move to a greater reliance on embodying its knowledge in output. *If* firms differ in their *ability* to embody innovation knowledge in their output, they will also clearly differ in their ability to benefit from a given innovation and therefore in their economically rational willingness to invest the resources required to innovate. This, in turn, will allow us to predict the functional locus of innovation when and if the differences in ability to appropriate benefit from output-embodied innovation knowledge are large enough to be observable under real-world conditions.

Whether or not and to what degree each of these conditions does in fact accurately describe the real world is an empirical matter which I will take up in the following sections of this paper. A simple example of the predictive power regarding the locus of innovation which we will acquire where these conditions do hold, however, can be seen in the following: Given case 2 conditions an independent inventor is much less likely to invent than are would-be innovators with other functional relationships to the innovation opportunity.

[2] If the above-described inability to predict the locus of innovation under case 1 conditions is to hold, costless enforcement of property rights is required for the following reason: since marketing of an innovation and enforcement of payment can be reasonably assumed to be costless for an innovating firm when it captures output-embodied benefit by utilizing the innovation knowledge in its own processes and/or products, *non*-costless marketing of an enforcement of payments for use of innovation knowledge by other firms would create a differential between benefit attainable from in-house and external use of the innovation and generate a preference for the former. This in turn would allow an incremental benefit from the same innovation to accrue to those innovators with a larger in-house use for it – and create a differential incentive to innovate as a function of locus of innovation.

[3] Suppose, for example, that a minor cost-reducing process innovation were made available to one of several manufacturers of a commodity with previously equal manufacturing costs, financial resources, etc. If further innovations or other changes did not intervene, the commodity producer benefiting from the innovation could in principle increase his market share as a consequence of innovation and thus "leverage" the direct benefits of the innovation, perhaps manyfold. But note that, even under such a set of circumstances, the innovator has no incentive to prefer to increase or decrease the market share of his own company relative to that of his competitors because he can, given perfect information, also charge the benefiting company for such second (and nth) order benefits arising from the innovation up to the point of indifference.

because an independent inventor has *only* non-embodied knowledge to sell.

3. Real-world ability of innovators to appropriate benefit from non-embodied innovation knowledge

In section 2 it was concluded that, if we were to be able to model the functional locus of innovation as a function of the appropriability of innovation benefit, innovators should *not* be able to effectively capture benefit from the licensing or sale of non-embodied knowledge regarding their innovation to others. Only two benefit capture mechanisms currently exist in the United States which allow innovators the possibility of capturing benefit from non-embodied innovation knowledge: (1) patent legislation (federal) allows an innovator to charge others for using freely available information published in his patent; and (2) trade secret legislation (state) allows an innovator to license knowledge to a user(s) and put the recipient under the legal duty of maintaining the secrecy of that information so that it will not become a free good on the marketplace. (Both of these mechanisms can also be used to capture benefit from output-embodied innovation knowledge, and we explore their effectiveness in this regard in section 4.2 below.)

3.1. Patent legislation as a mechanism for capturing benefit from non-embodied innovation knowledge

A patent grants an inventor the right to exclude others from the use of his invention for a limited period. In return for the right to exclude not only those who copy the invention but also those who independently discover the same thing, the inventor must disclose the invention to the public at the time of the patent's issue. This disclosure, contained in the patent itself, must be sufficiently detailed so that those "ordinarily skilled in the art" may copy and utilize the invention after the patent's expiration. While considerable information exists on the number of patents acquired by various firms and industries over time and on the various correlations between such "patent rates", firm size, R&D expenditures, and similar variables, very little information exists on the real-world effect of a patent grant on an inventor's ability to gain benefit from the non-embodied knowledge characterizing his invention [13]. I review the available empirical data below.

Evidence of a patent system's effectiveness as a mechanism for allowing the capture of benefit from non-embodied innovation knowledge and/or benefit from output-embodied innovation knowledge can be seen in its influence on an innovator's willingness to invest in research and development, while evidence of its effectiveness in allowing benefit capture from non-embodied knowledge *only* can be seen via data on license agreements and related payments. A recent study by Taylor and Silbertson [14] provides both types of evidence. [4] Evidence regarding the effect of patent protection on an innovator's willingness to invest in R&D was obtained via a questionnaire ("Form B") which asked: "Approximately what proportion of your R&D in recent years would not have been carried out if you had not been able to patent any resulting discoveries?" [16]. The data derived from this question are shown in table 2. Note that 24 of the 32 returns indicate that only 5% or less of recent R&D expenditures would not have been undertaken if patent protection had not been available [17].

A direct measure of the ability to capture benefit from non-embodied innovation knowledge afforded to innovators by patents may be obtained

[4] Taylor and Silberston examined the impact of British and foreign patents on a sample of 44 British and multinational firms involved in five broad "classes" of industrial activity: chemicals (including pharmaceuticals and petrochemicals); oil refining; electrical engineering (including electronics); mechanical engineering; and man-made fibers. Approximately 150 firms were invited to join the study. Coded as being in one of the five specified classes, they were selected from a "comprehensive list of U.K. quoted companies" on the basis of their net assets in 1960: In each class all companies showing net assests in excess of 10 million pounds in 1960 were selected, and every seventh company of the remainder was selected from a list tabulated in ascending order of net assets in 1960. Finally, "some additions were made to take account of mergers and acquisitions and to include unquoted companies". Eventually "just over 100" firms responded to the letter of invitation. Sixty-five expressed interest, but "some twenty of these indicated that patents were a very minor aspect of their operations and were firmly believed to have no significance on the business... this left 44 firms which agreed to participate in the inquiry" [15]. Of these, 30 ultimately agreed to participate fully and fill out the detailed questionnaires provided by the authors, while the remaining 14 agreed to provide more limited information and to be interviewed.

Table 2
Estimated proportions of R&D expenditure dependent on patent protection: twenty-seven responding companies [a]

Industry	Estimate of R&D affected [b]				
	None or negligible	Very little (less than 5%)	Some (5–20%)	Substantial (over 20%)	Total returns
	Number of returns				
Chemicals:					
– Finished and speciality	1	2	1	4	8
– Basic	1	2	1	0	4
Total chemicals	2	4	2	4	12
Mechanical engineering	7	1	0	2	10
Man-made fibers	1	1	0	0	2
Electrical engineering	7	1	0	0	8
Total	17	7	2	6	32 [c]
Percentage of returns	53%	22%	6%	19%	100%

[a] Table redrawn from Taylor and Silberston [22, table 9.1, p. 107].
[b] Percentages refer to the estimated reduction in annual R&D expenditure in recent years that would have been experienced, had patent monopolies not been available.
[c] Some companies made returns for more than one activity.

by looking at licensing cost and benefit data. To the extent that an effective patent monopoly is provided to an innovator, he might choose to exercise it by a policy: (1) excluding all competitors; (2) selectively licensing some applicants; or (3) licensing all applicants for a royalty and/or other consideration. If the innovator chooses to reap benefit from non-embodied innovation knowledge via his patent monopoly by use of policy option (3), licensing all comers, diffusion of the innovation may be assumed freely to occur and the maximum value of benefit from non-embodied knowledge capturable by the innovator via the patent mechanism can be approximately represented by licensing fees and/or other considerations received minus patenting and licensing costs incurred by the innovating firm. In the event, most firms studied by Taylor and Silberston claimed to be following policy option (3), a policy of licensing all "responsible" applicants, rather than options (1) or (2). Indeed, the authors note, "we were repeatedly assured that the main problem for the licensing department is to interest reputable firms in taking licenses rather than dissuading them from doing so, and many licensing specialists to whom we talked were plainly puzzled that their task might be seen in the latter rather than the former light" [18]. Patent-related cost and benefit data provided by Taylor and Silberston's "main sample" of 30 firms will be found summarized in table 3.

Taken together, tables 2 and 3 suggest that, except in the pharmaceutical field (for particular reasons noted in footnote 9 below), firms do not find the patent grant to be of significant benefit.[5]

[5] A study performed by a group of candidates for the Master's Degree at Harvard Business School [19] also contains some information on the value of patents to firms which hold them. A questionnaire was pilot tested, modified, and then sent out to a sample of 266 firms known to hold a relatively large number of patents [20]. Sixty-nine of the questionnaires (26%) were completed and returned in time to be included in the study's analysis phase. All but four of these respondents held more than 100 patents and collectively they "held approximately 45,500 patents, or about 13.5% of all the unexpired U.S. patents held by domestic corporations at the end of 1956" [21]. One of the questions attempted to determine the importance of patents to firms by asking the "executive responsible for technical change" to "please state briefly the importance of patents to the company". Thirty-seven responded in a manner which the students felt they would clearly categorize as follows: "very important", 8; "some importance", 14; "not very important", 15 [22]. While, unfortunately, neither the question nor the coding categories used are clear on what interviewers or interviewees meant by "important", we find the results suggestive in light of the Taylor and Silberston data: 40% of a sample of interviewees from companies selected because they patent a great deal felt that patents are "not very important to their companies".

Table 3
Relationship of 1968 patent expenditures to 1968 patent-related receipts in Taylor and Silberston "main sample" of thirty companies [14]

Industry	1 1968 UK license and royalty receipts [a] £ (million)	2 1968 UK patenting and licensing expenditures [b] £ (million)	3 1968 R&D expenditures in UK [c] £ (million)	4 1968 license receipts as % of R&D expenditures plus patenting and licensing expenditures (cols. 1 – [2 + 3])	5 1968 license receipts as % of 1968 UK sales col. 1 / note d
Chemicals					
– Pharmaceuticals	3.7	NA	7.1	NA	6
– Other finished and speciality	0.2	NA	10.1		0.04
– Basic	2.4	NA	3.3	NA	1
Total chemicals	6.3	0.99	20.5	29	1.1
Mechanical engineering	1.4		7.3	18	0.4
Man-made fibers	0.7	0.37	7.6	9	0.2
Electrical engineering	2.3	0.65	50.5	4	0.3

Except as noted in a–d below, data in all columns were derived from the same set of companies. N.B. that Taylor and Silberston have *not* logged patent and R&D expenditures data relative to receipt data on licensing, royalty, and sales. *All* table 2 data are for 1968.

[a] *Source:* Taylor and Silberston, table 8.7, p. 164. (T&S note that data from oil companies in sample and "one large electrical" group are excluded from table 8.7.)

[b] *Source*: Taylor and Silberston, table 6.4, p. 109. (T&S note that data from oil companies are excluded from table 6.4.)

[c] *Source*: Taylor and Silberston, table 8.1, col. 2, p. 10?. I have excluded oil company data from basic chemical category to make this data base more compatible with table 6.4. T&S offer more aggregated R&D expenditure data in table 6.4, whose magnitudes deviate from those shown in table 8.1 by 20–40%. These discrepancies are unexplained, but our uses of that data are not sensitive to corrections of this magnitude.

[d] *Source*: Taylor and Silberston, table 8.1, col. 4, p. 145.

This finding has emerged in the face of three study elements which would tend to raise the level of benefit shown: (1) the authors noted in their discussion of sample selection (see footnote 4 outlining the study methodology) that firms which did not feel that patents significantly affected them tended to decline to join the study sample; (2) the authors noted that, "to avoid understating the impacts of patents", they chose to "err on the high side" [18] in their acquisition of data for table 1; (3) the authors also noted that the license agreements which resulted in the costs and benefits shown in table 2 involved the transfer of and payment for valuable unpatented "know-how" in addition to the transfer of information protected by patents and that "this may result in some overstatement of the true payment for patent licenses themselves". Note, however, that some understatement of real benefits may also be present because remissions of any *non*-monetary benefits (e.g. cross-licensing) are omitted from table 2 [23].

Another study whose data can be used to assess the possible benefits from non-embodied knowledge that corporations reap through licensing of their patents was conducted by Wilson [24] who reports data on royalty payments submitted by some U.S. corporations to the U.S. Securities and Exchange Commission in 1971 on Form 10K.[6]

[6] In 1971 firms wre required to report royalty payments if they were "material" with the precise interpretation of that term being left up to individual firms. Focusing on the *Fortune* listing of the 1000 largest manufacturing corporations in 1971; Wilson found that 518 had considered their royalty receipts "material" enough to report to the SEC. Since he was interested only in royalty payments for "technology licenses", he used various means to detect and winnow from the sample firms which reported royalty payments for such things as trademarks, copyrights, and mineral rights [25]. The end result of this process was a sample of 350 royalty figures for 1971 which Wilson felt were largely or entirely payments

Table 4
Wilson and Taylor-Silberston royalty payment data compared

Industry	Wilson [24] (1971 US data)		Taylor and Silberston [14] (1968 UK data)	
	% of US sales by firms in sample [a]	Royalties paid as % of firm 1971 sales [a]	Royalties paid as % of firm 1968 sales [b]	"Industrial activity"
Chemicals				Chemicals
- Industrial	76.4	0.244	0.042	- Basic
- Drug	72.8	0.745	0.635	- Pharmaceuticals
- Other	51.4	0.034	0.044	- Other finished and speciality
Machinery	40.2	0.051	0.255	Mechanical engineering
Electrical	40.5	0.13	0.182	Electrical engineering

[a] *Source*: Wilson [24, table 12, p. 169]. Note that the data presented here are computed from Wilson's sample of 350 royalty reports, *not* his larger sample comprised of these reports plus estimated data.
[b] *Source*: Royalty and license fee expenditures data from Taylor and Silberston [14], table 8.7, col. 3, p. 164, sales data from table 8.1, col. 4, p. 145. (Petrochemicals have been removed from the basic chemicals category of table 8.1 to make this category compatble with the equivalent category of table 8.7.)

The reader will find Wilson's data for the SIC categories apparently most similar to the "industrial activity classes" examined by Taylor and Silberston compared in table 4.

Even though derived from a different source and country, the Wilson data have magnitudes quite similar to the Taylor and Silberston data. While unfortunately the table 4 data are for royalty payments rather than receipts (the Wilson data providing information on payments only), it is likely that the bulk of technical agreements would be between firms in the same industry.[7] If so, it would follow that the low magnitude of royalty payments in the Wilson data implies that royalty receipts would also be found low in the industries sampled. This would be in line with the Taylor and Silberston data indicating that the benefit captured by innovators from the sale of non-embodied knowledge is indeed low in most industries.

The slim data base I have just reviewed indicates that, in industry aggregate terms, innovators do not capture much benefit from the sale of non-embodied innovation knowledge via the patent mechanism. Are these data congruent with "tests of reason" which one can apply to the matter? Let us explore. First, does it make economic sense that firms would take out patents if these do not, on average, yield much economic benefit? The answer is yes – because the cost of applying for patents is also low. The cost of the average patent application prosecuted by a corporation is on the order of \$5,000 today.[8] (Even this small cost is often not very visible to corporate personnel deciding on a patent application "purchase" because it is typically subsumed within the overall cost of operating a corporate patent department.)

for "technical agreements", a term he does not define, but which presumably includes both patent and technical know-how-related payments. The responses of these 350 firms were then aggregated under appropriate "2 and 3 digit SIC codes" (not given) and displayed in tabular form. (Wilson used the 350 reports of corporate royalty payments to develop estimates of royalty payments to all members of the industries he studied, and then compared these estimates with industry-level data on corporate R&D expenditures collected by the National Science Foundation. As I find Wilson's estimating procedures inappropriate for our purposes here, I use only the direct company report data he provides.)

[7] This point is never explicitly examined, but is apparently assumed in Taylor and Silberston [14]. See especially the in-depth studies of Pharmaceuticals, Basic Chemicals, and Electronics in that source.

[8] In 1961 the Commissioner of Patents reported the cost of an average patent application prosecuted by a corporation to be \$1,000 to \$2,500, and the cost of a single application prosecuted by an attorney for an individual to be \$680 [26]. My own recent conversations with several corporate patent attorneys yielded an estimate that the "average patent application prosecuted by a corporation" currently costs on the order of \$5,000.

Second, what do we know about the nature of the patent grant and of the real-world workings of the patent office and the courts? And, is it reasonable in the light of what is known to conclude that the patent grant is likely to offer little benefit to its holder? Consider the following three points.

(1) It is important to note that a patent, if valid, gives a patentee the right to exclude others from using his invention, but it does *not* give him the right to use it himself if such a use would infringe the patents of others. For example, Fairchild has a patent on the so-called planar process, an important process invention used in the manufacture of integrated circuits. If firm B invents and patents an improvement on that process, it may not use its improvement invention without licensing the planar process from Fairchild and Fairchild may not use the improvement either without licensing it from firm B. Thus, in rapidly developing technologies where many patents have been issued and have not yet expired, it is likely that any new patent cannot be exercised without infringing the claims of numerous other extant patents. Given this eventuality, the benefit of a particular patent to an inventor would very probably be diminished because he might be prevented from using his own invention or he might be forced to cross-license competitors holding related patents in order to practice his invention.

(2) The patent system places the burden on the patentee of detecting an infringer and suing for redress. Such suits are notoriously long and expensive and both defendants and plaintiffs tend to avoid them assiduously. For the defendant the best outcome in recompense for all his time and expense is judicial sanction to continue this alleged infringement, while the worst outcome would involve the payment of possibly considerable penalties. For the plaintiff the likelihood that a court will hold a patent valid and infringed – as opposed to invalid and/or not infringed – is on the order of one to three [27]. If a patentee has licensees already signed up for a patent at issue, he has a high incentive to avoid litigation: If he loses, and the odds are that he will, he loses payments from all licensees, not just the potential payments from the particular infringer sued.

(3) The patent grant covers a particular means of achieving a given end but not the end itself, even if the end and perhaps the market it identifies are also novel. A would-be imitator can "invent around" a patent if he can invent a means not specified in the original inventor's patent. In the instance of the Polaroid and Xerography processes and a few other notable cases, determined competitors could not, in fact, invent around the means patented by the inventor. In most instances and in most fields, however, inventing around is relatively easy because there are many known means by which one might achieve an effect equivalent to the patented one, given the incentive to do so. Where inventing around is possible, the practical effect is to make the *upper* bound value of an inventor's patent grant equal to the estimated cost to a potential licensee of such inventing around.

Taken in combination, the observations made above may be applied to provide a very reasonable explanation for the relatively low benefit from non-embodied knowledge which we have found innovators in most fields obtaining via the patent grant.[9] Thus, in sum, we see via both data and test

[9] As an example, consider the application of these observations to the value of patents obtained in the field of semiconductor electronics.

The semiconductor field is currently a very fast-moving one in which many unexpired patents exist which address closely related subject matter. The possible consequence – confirmed as actual by corporate patent attorneys for several US semiconductor firms whom I interviewed – is that many patentees are unable to use their own inventions without the likelihood of infringing the patents of others. Since patents challenged in court are unlikely to be held valid, the result of the high likelihood of infringement accompanying use of one's own patented – or unpatented – technology is not paralysis of the field. Rather, firms will in most instances simply ignore the possibility that their activities might be infringing the patents of others. The result is what Taylor and Silberston's interviewees in the electronic components field termed "a jungle", and what one of my interviewees termed a "Mexican Standoff". Firm A's corporate patent department will wait to be notified by attorneys from firm B that it is suspected that A's activities are infringing B's patents. Since possibly germane patents and their associated claims are so numerous, it is in practice usually impossible for firm A – or firm B – to evaluate firm B's claims on their merits. Firm A therefore responds – and this is the true defensive value of patents in the industry – by sending firm B copies of "a pound or two" of its possible germane patents with the suggestion that, while it is quite sure it is not infringing B, its examination shows that B is in fact probably infringing A. The usual result is cross-licensing with a modest fee possibly being paid by one side or the other. Who pays, it is important to note, is determined at least as much by the contenders' relative willingness to pay to avoid the expense and bother of a court fight as it is by the merits of the particular case.

of reason that the patent grant does not effectively enable innovators to capture benefit from non-embodied innovation knowledge in most fields.

3.2. Trade secret legislation as a mechanism for capturing benefit from non-embodied innovation knowledge

Trade secrets, like patents, can be used to capture benefit from non-embodied innovation knowledge. (As noted earlier, their effectiveness in capturing benefit from output-embodied innovation knowledge will be explored in section 4.2 below.) Trade secrets, also sometimes termed "know-how", typically refer to inventions and/or knowledge which can be kept secret even *after* development is completed and commercial exploitation begun. The possessor of a trade secret has an indefinite period of exclusive use of his invention or discovery. Trade secret legislation allows him to keep the information entirely secret or to make legally binding contracts with others in which the secret is revealed in exchange for a fee or other consideration and a commitment to keep the information secret. A trade secret possessor may take legal steps to prevent its use by others *if* they can be shown to have discovered the secret through unfair and dishonest means such as theft or breach of a contract promising to keep it secret.

A legally protectable monopoly of indefinite duration would appear to make trade secrecy a very attractive mechanism for capturing innovation benefit. It is, however, an option only for innovations which can in fact be kept secret since the holder of a trade secret cannot exclude anyone who independently discovers it or who legally acquires the secret by such means as accidental disclosure or "reverse engineering". In practice, trade secrets have proven to be effective only with regard to product innovations incorporating various technological barriers to analysis, or with regard to process innovations which can be hidden from public view.

There are, in the first instance, certain innovations embodied in products which, while sold in the open market and thus available for detailed inspection by would-be imitators, manage nevertheless to defy analysis for some technological reason and which cannot therefore be reverse engineered. Complex chemical formulations sometimes fall into this category, the classic case being the formula for Coca-Cola. Such barriers to analysis need not be inherent in the product – they can sometimes be added on by design. Thus, some electronic products gain some protection from analysis via use of a packaging method ("potting") and packaging materials which cannot easily be removed without destroying the proprietary circuit contained within [28]. Methods for protecting trade secrets embodied in products accessible to competitors need not be foolproof to be effective – they

Thus, in the semiconductor field, except for a very few patent packages which have been litigated, which have been held valid, and which most firms license without protest – notably the Bell transistor patents and the Fairchild planar process patents – the patent grant is worth very little to inventors who obtain it. Indeed, the one value suggested to us – defense against the infringement suits of others – suggests that perhaps the true net value of the patent system to firms in the semiconductor industry is negative because it requires all to assume the overhead burden of defensive patenting.

In sharp contrast to the situation pertaining in most other industries and the electronics field in particular, the patent grant seems to confer significant benefit to innovators in the pharmaceutical field, as indicated by the Taylor–Silberston and Wilson data discussed in tables 3 and 4. My own discussions with corporate patent attorneys working for pharmaceutical firms brought out two likely reasons: (1) unusually "strong" patents are obtainable in the chemical field, of which pharmaceuticals is a part, and (2) it is often difficult to "invent around" a pharmaceutical patent. Pharmaceutical patents can be unusually strong because one may patent an actual molecule found to have useful medical properties and its analogs (in contrast to only the *particular* means to a given end in other fields). One need not make each analog claimed, but can simply refer to lists of recognized functional equivalents for each component of the molecule at issue. For example, if a molecule has ten important component parts, one patent application might claim X plus 10 recognized functional equivalents of X for each part. Obviously, by this means an inventor may claim millions of specific molecules without actually having to synthesize more than a few. Furthermore, demonstration that any of the analogs so claimed does not display the medical properties claimed does not invalidate the patent.

Pharmaceutical patents are difficult to "invent around" because the mechanisms by which pharmaceuticals achieve their medical effects are usually not well understood. Thus, would-be imitators do not gain much insight by examining a competitor's patented molecule proven to produce a desired medical effort. Eloquent testimony to this fact is provided by the pharmaceutical industry's research practice of synthesizing great numbers of molecules and "screening" these for possible medical activity rather than synthesizing only a few molecules predicted to have a given activity. The knowledge required for such prediction is seldom available today.

simply have to raise enough of a barrier in a given case to create an unattractive cost benefit equation for would-be imitators in that case.

In the second instance, process innovations such as novel catalysts or process equipment can be protected effectively as trade secrets, whether or not they could be "reverse engineered" by a would-be imitator allowed to examine them, simply because they can be exploited commercially while shielded from such examination behind factory walls.

Few empirical data exist on the information protected as trade secrets: There is no central registry for such material analogous to the U.S. Patent Office, and even those trade secrets which are revealed to others to obtain benefits from non-embodied innovation knowledge, the subset of interest to us here, are contained in private contracts which do not usually appear on any public record unless litigated [29]. While some examples exist of major benefits from non-embodied knowledge being reaped by innovators via licensing of trade secrets [30], I argue that the typical effectiveness of this mechanism is severely limited for two reasons. First, the mechanism is clearly not applicable to product or process innovations which are not commercially exploitable while concealed behind factory walls and which are amenable to reverse engineering if accessible to inspection by imitators – and these considerations apply to many industries and many innovations. Second, a trade secret licensor can only gain redress under trade secret legislation if he can document the *specific* illegal act which diffused his innovation to unlicensed parties. A licensor finds such specificity difficult to achieve if he seeks to license non-embodied knowledge to many licensees.

4. Real-world appropriability of benefit from output-embodied innovation knowledge

In the previous section we found that an innovator's ability to appropriate benefit from non-embodied knowledge is low in most industries. If this is so, then significant economic reward, if any, must come primarily from the innovator's ability to appropriate benefit from output-embodied knowledge. The logical necessity of this conclusion is clear – the two categories of economic benefit are mutually exclusive and jointly exhaustive.

The ability of an innovator or innovating firm to capture benefit from output-embodied innovation knowledge derives from its ability to establish a quasi-monopoly position with respect to that innovation. I propose that two "levels" of quasi-monopoly are germane: (1) quasi-monopoly which an innovation affords to the entire industry of which the innovator is a member, and a portion of which the innovator derives in accordance to his "size"; and (2) quasi-monopoly which an innovation affords to the single innovating firm relative to other members of his industry. The ability to capture benefit from output-embodied innovation knowledge which these two levels of quasi-monopoly afford to firms is additive. I examine each, and the mechanisms by which each is achieved. While related empirical data are also explored in this section, I have found it to be so sparse on the issues addressed that the findings can best be seen as suggestive. Research approaches discussed, on the other hand, offer useful models for the additional empirical work required.

4.1. *Benefit from output-embodied knowledge appropriable by an innovating firm via creation of an industrywide quasi-monopoly*

I define an industry as made up of all firms making products which are close substitutes (i.e. have high cross-elasticity of demand). Firms in an industry may share in an industrywide quasi-monopoly if significant barriers exist which deter free entry to the industry by additional firms. Examples of such barriers to industry entry are specialized facilities, specialized production skills, and specialized sales forces, which are required for functioning effectively in an industry, which are possessed by firms already in that industry, but which must be acquired by potential new entrants.

Barriers to industry entry by new firms are common but difficult to measure. Consider, as an example, the barriers which face a firm which is a member of an industry characterized by a given functional relationship to an innovation (e.g. an industry which uses semiconductor process equipment to make semiconductors) and which wishes to join an industry characterized by another functional relationship to that innovation (e.g. the industry which manufactures semiconductor process equipment).[10] These two types of firms are really in very different businesses. Each has a great deal of know-how, organizational arrangements, and capital equipment which is quite specialized to build its existing products and to serve its existing customer base. Thus, the semiconductor manufacturer has a sales force which specializes in serving semiconductor buyers. This force would be entirely inappropriate for selling semiconductor process equipment: the customers are different, the sales techniques are different (samples of semiconductor devices can be given out as a selling technique, but not samples of semiconductor process equipment), and the specialized knowledge which the salesman must have is completely different (a salesman with an electrical engineering background can help customers with problems in selecting and using semiconductor devices; a background in solid state physics would be considerably more appropriate for a salesman trying to sell the semiconductor process equipment used to grow the ultrapure single silicon crystals used in semiconductor device manufacture).

If the sales, organizational, and production infrastructure which a company uses to serve one functional role relationship to a given innovation cannot effectively be used in the service of a different functional relationship, then it follows that a firm wishing to change such relationships must also set up a new infrastructure appropriate to this new role. Further, since the costs of the infrastructures of competitors already having the role relationships the innovator wishes to acquire are typically allocated across many products (e.g. a "line" of process equipment or a "line" of semiconductor devices), the would-be new entrant must develop/adopt/buy a similar line of product to sell if he wishes to be economically competitive. All these requirements, I suggest, represent significant barriers to industry entry.[11]

Where significant barriers to industry entry do exist, an innovation made by one member of the industry can establish an industry-level quasi-monopoly with respect to that innovation which in turn can allow the industry as a whole to increase its rate of profit and/or volume of sales and thus reap benefit from output-embodied innovation knowledge. As an example, consider an innovation in plastics molding machinery made by a producer of a commodity plastic such as polyethylene. Assume the innovation allows molders of plastic items to significantly decrease their production costs. Further assume, as is realistic, that machinery innovation itself cannot be protected effectively via patent or other means by the innovator and that the machine works equally well using polyethylene manufactured by any supplier of such. Under these circumstances adoption of the innovation by molders might well increase demand for polyethylene more rapidly than supply could respond (it takes many years to build a new polyethylene plant) and the profits of all polyethylene producers – molding innovation developer and other producers alike – would then rise in proportion to their market share for polyethylene.

The assumptions embedded in the above machine innovation example – that the innovator has no ability to control or benefit from the diffusion of non-embodied knowledge regarding his innovation, and that the innovation benefit is instantly distributed to all competitors currently in the industry (i.e. increased profits on polyethylene are afforded to the innovating and non-innovating polyethylene producers simultaneously) – are equivalent to assuming the innovation to be a privately financed collective good. This being the case, the argument developed by Mancur Olson in his *The Logic of Collective Action* [31] can be applied to predict that the firm with the most to gain from the innovation is the one most likely to

[10] Note that firms holding different functional relationships to a given innovation are indeed in different industries according to the definition of "industry" cited previously, and that it is important to our model that barriers to entry exist between these industries. This is so because if it were easy, for example, for an innovation product manufacturer to become a product user at a moment's notice should such a course of action seem to promise an increased ability to capture benefit from the innovation, we would only be able to predict the functional locus of innovation in a weak sense, i.e. "the developer of X innovation will *become* a user" rather than able to make the stronger statement that "the developer of x innovation will be a firm and/or individual which currently is a member of the user community".

[11] It is important to note, however, that barriers to entry to a new industry (barriers to adding a new functional role with respect to a given innovation) may be considerably reduced if a firm does *not* want to make a full-scale entry into a new industry but simply wants to vertically integrate and *only* supply its own needs. Thus, if a semiconductor process machine user wishes to build a few units of an innovative process machine for in-house use, it does not need a sales force, an external field service force, nor a full line of equipment in order to spread the cost of these.

provide it to the group. (Firms holding any functional relationship to a given innovation are group members in Olson's sense if their relationship allows them the possibility of deriving output-embodied benefit from the innovation. The qualitative nature of the output in which the innovation benefit is embodied will differ, of course, in accordance with the functional role relationship of group member to innovation. For example, if the process equipment at issue is a plastics molding machine capable of making parts more cheaply, an equipment manufacturer's benefit is embodied in sales of the innovative molding machine; a plastic supplier's benefit is embodied in increased sales of plastic molding material.)

When and if industry-level quasi-monopolies do indeed provide significant benefits to would-be innovators, we should be able to observe empirically a concentration of innovations among what Olson terms the "larger" group members (quasi-monopoly participants). At the moment the only study I am aware of which offers an empirical research model that could test this hypothesis is by von Hippel [32]. [12] The study focuses on semiconductor process machinery innovations developed by firms that use such machinery in the manufacture of silicon-based semiconductors and contains data on the market share ranking of innovating user firms [13] in the year in which their sampled process machinery innovations were first used for commercial production of silicon-based semiconductors. This market share data can serve as an approximate measure of the relative amount of benefit from output-embodied innovation knowledge potentially appropriable by members of the sampled group of innovating user firms if we assume, as previously noted, that an innovation, once made by any one group member, becomes a collective good instantly provided to all members of that group. (Given this assumption, it is reasonable to conclude that the pre-innovation market shares of all group members whose outputs embody the innovation benefit will remain constant post-innovation. And if this is so, we may usefully approximate group member size by a group member's market share of silicon-based semiconductors at the time of the innovation's first commercial use.)

Note from table 5 that, four out of the five innovating user firms identified are ranked among the largest eight firms in terms of share of market in the year of first commercial use of their innovation(s). [14] This is the result we would expect if a significant industry-level quasi-monopoly existed and Olson's hypothesis were correct. I emphasize, however, that the results of this single study can only be seen as suggestive as it does not address reasonable alternative explanations for the finding (for example, it offers no information on the direction of causality involved in the observed correla-

[12] The method by which market share data were acquired in the study is fairly straightforward and is summarized in the notes to table 5. The method by which the sample of process machinery innovations was selected involved, first, selecting a subset of all process steps involved in each type of manufacture of silicon-based semiconductors. (Process steps and innovations studied are explicity identified in von Hippel [33], table 1.) For each process step selected, the process machinery (if any) used in the initial commercial practice of that step was identified and included in the sample. Next, all subsequent improvements to process machinery for each step which offered a major improvement in functional utility to the user of such machinery (judged relative to previous best practice used in commercial manufacture) were identified and added to the sample. Finally an exhaustive list of process machinery innovations which offered any increment in functional utility to the user was collected for one randomly selected process step and these made up a sample of minor improvement innovations. All process equipment innovations in the sample were successful in the sense of receiving widespread use in their respective industries and becoming a commercially viable industrial good manufactured for commercial sale by at least one (and usually several) process equipment firms.
The "source" of each sampled innovation was determined via literature searches and interviews with user and manufacturer personnel. An innovation "source" was the firm which developed and built the first unit of equipment embodying the innovation which was used to produce commercially sold semiconductors. Innovations found to have a user source were coded as shown in table 5.

[13] Only firms with a use relationship to the sampled innovations are included in this study. Would-be innovators bearing other functional relationships to those innovations such as semiconductor machinery manufacturers, while also clearly in a position to gain benefit from output-embodied innovation knowledge and thus group members in Olson's sense, are excluded. This exclusion has no practical consequence here since, for reasons analogous to those spelled out in section 5 for the pultrusion industry study, it is quite certain that the "largest" group members with respect to this innovation sample are innovation users.

[14] Firm coded NA in table 5 were *not* smaller firms than those specifically identified: rather, in these instances, several major firms moved on the innovation so rapidly that I was unable accurately to determine retrospectively which of these had priority.

Table 5
Market rank of innovating vs. non-innovating user firms

First innovating user firm	1 Number of process equipment innovations [a]			2 Date of first commercial use of process equipment innovation [a]	In year of first commercial use of process innovation			
					3a Innovative firm's semiconductor shipments [b]	3b	4 Number of US semiconductor firms extant [d]	5 Sales of parent firm [e] (total) $ (million)
	Initial	major	minor		$ (million)	Industry rank		
Fairchild	1	1	3	1959	20	6	34	43
				1960	27	6	47	68
				1966 (3)	146	2	50	207
IBM		2	1	1965		5 [c]	50	3,700
				1965		5 [c]	50	3,700
				1967		5 [c]	53	5,300
Western Electric	1	1		1956	4.5	6	26	
				1960	27	6	47	7,900
Hughes	1			1970		NA	NA	NA
Motorola		1		1961	28	6	53	298
NA	2	5	1					

[a] Data from von Hippel [33] and see text.
[b] Share of market rankings are derived by conversion of Tilton [34] data (p. 66, table 4–5) on percent of semiconductor shipments attributable to major firms into rankings (shipments data include in-house and government sales). Firms with the same shipment percent in given year are all given the same rank. Tilton's share of market data only covers the years 1957, 1960, 1963, and 1966. For innovations whose date of first commercial use (col. 1) falls between these years, data on the nearest of the years examined by Tilton are used. Conversion of SOM rankings into $ shipments was effected via use of Tilton's data on total semiconductor shipments (p. 90, table 4–7).
[c] IBM has, since 1962, been a major producer of silicon semiconductors for in-house use only, and thus "shipments" data are not available to determine IBM's market share rankings. Industry "guesstimates" of IBM's ranking in 1965 and 1967 place that firm conservatively among the top five producers for those years.
[d] Tilton, p. 52, table 4–1.
[e] Data from annual reports of parent companies. Fairchild was acquired by Fairchild Camera and Instrument in 1954, and therefore sales figures of the parent company are shown.

tion between innovation rates and market share). More empirical work will clearly be required on the issue.

4.2. Benefit from output-embodied innovation knowledge appropriable by an innovating firm via creation of a firm-level quasi-monopoly

I now move to a consideration of the mechanisms by which an innovating firm might hope to establish a quasi-monopoly with respect to all other firms, both current competitors and those currently outside the industry, and thus be in a position to capture benefit from output-embodied innovation knowledge via increases in profit rates and/or sales volume. I suggest that there are only three such mechanisms extant – patents, trade secrets (also termed know-how) and "response time" – when we exclude from consideration those comparative advantages one firm may have over another which, while they may aid an innovator, are really innovation-independent and may equally serve an imitator (e.g. a relatively favorable position with regard to finances, mineral rights, marketing channels, firm reputation, etc.). We will discuss each of these three extant mechanisms below.

As noted in section 3.1, a patent grants an inventor the right to exclude others from using his invention for a limited period in exchange for public disclosure of that invention. Patent legislation requires that this public disclosure be made at the time of the patent's issue and be in sufficient detail so that others "ordinarily skilled in the art" may readily imitate the invention, presumably upon the patent's expiration. The result of the public disclosure is that interested imitators have access to the invention and must be constrained by law, rather than by lack of knowledge, from using it – *if* the inventor is to be able to use the patent grant as a mechanism for maintaining a quasi-monopoly and garnering benefit from output-embodied innovation knowledge via his own exploitation of the invention. But, as we have seen previously, the law offers little effective protection to patent holders. The burden of finding any infringement is on the patentee – no mean task, particularly if the infringement does not involve a product sold on the open market but rather a process or machinery invention which an infringer may exploit and benefit from in the privacy of his factory. Moreover, the burden of prosecuting the infringer also falls on the patentee. Such prosecutions are notoriously long and expensive and studies of court records [27] have shown that the likelihood of a patent being held valid and infringed are on the order of three to one against the patent holder. Thus, the same evidence that led me to conclude earlier that the patent grant was not an effective mechanism for the capture of benefit from non-embodied innovation knowledge also leads me to conclude that the patent grant is not an effective mechanism for the capture of benefit from output-embodied innovation knowlege.

Trade secrets, a second possible means for the establishment of innovation-based quasi-monopolies at the level of the firm, refer to innovations which can be kept secret *after* development is completed and commercial exploitation begun. As was explained in section 3.2, secrecy can be maintained during commercial exploitation either because (1) the innovation cannot be "reverse engineered" and imitated even though available to inspection by would-be imitators skilled in the relevant analytical tools (the formula for Coca-Cola is the classic example of such) or (2) the innovation, while susceptible to reverse engineering if opened to the inspection of would-be imitators, can be hidden from such inspection by some means (e.g. process equipment developed by users and shielded within their own firms). As was noted earlier, essentially no hard data exist on the effectiveness with which innovations kept as trade secrets allow firms to establish firm-level quasi-monopolies and capture benefit from output-embodied innovation knowledge. I am aware of two types of anecdotal data, however, which suggest that trade secrets can sometimes be a very effective benefit capture mechanism. First, many whom I have interviewed in corporations feel that the mechanism is very effective for innovations which can in fact be kept secret. (Logically, it is likely that the trade secret mechanism will be more effective in allowing the capture of benefit from output-embodied innovation knowledge than in allowing the capture of benefit from non-embodied innovation knowledge, as the latter use requires diffusion of the secret beyond the confines of the innovator's factory while the former does not.) Second, in some industries one can observe that firms incur significant expense to insure that outsiders do not get the chance to inspect their production equipment and/or techniques – implying that these firms do regard the knowledge protected as having significant economic value. [15] Clearly, more research into the effectiveness of trade secrets would be valuable.

The third mechanism noted above is one I term "response time". I define it as the period an imitator requires to bring an imitative product to market or to bring an imitative process to commercial usefulness when he has full and free access to any germane trade secrets or patented knowledge in the possession of the innovator. Response time exists simply because many barriers in addition to lack of knowledge must be overcome in order to bring any product or process – even an imitative one – to commercial reality. Engineering tooling must be designed, materials and components ordered, manufacturing plants made ready, marketing plans developed, etc. During the response time period an innovator by definition has a monopoly and is in a position to capture benefit from output-embodied innovation knowledge by increasing his rate of profit and/or his market share.

[15] Interestingly, there is a wide variation in the amount of effort firms exert to prevent inspection of their process trade secrets. In some firms and industries access is denied even to repairmen wishing to repair standard equipment located near proprietary equipment. In other firms and industries I have observed a willingness to allow free inspection of proprietary equipment and even a willingness to encourage its commercial manufacture and sale by others [30]. Such objectively codable differences in behavior may prove useful as one research measure of the economic value of trade secrets.

In principle, if an imitator became aware of an innovator's protected knowledge at the moment he developed it there would be no response time protection for the innovator: both innovator and imitator could proceed with commercialization activities in tandem. Response time is an important innovation benefit capture mechanism in reality, however, because would-be imitators seldom become aware of an innovator's knowledge at the moment he develops it. Typically, in fact, an imitator only becomes aware of a promising new product when that product is introduced to the marketplace. Until that point the innovator has been able to protect his product from the eyes of interested competitors inside his factory. After that point, if the product is easily reverse engineered and has no patent protection only the response time mechanism can provide him with some quasi-monopoly protection from imitators.

The real world value of response time to innovators is suggested by the elaborate lengths to which innovators sometimes go to hide their new product plans prior to marketplace introduction – and the sometimes equally elaborate affects of would-be imitators to ferret these plans out. As in the case of trade secrets, however, little formal research exists on the value of response time – and what there is of it addresses "lead time" rather than response time.[16] Data from the one study I am aware of which touches on the correlation between the commercial success of a sample of industrial products and lead time, Project Sappho [36], indicate that the effect of lead time – and its response time component – may sometimes be easily discernable, however. If so, these could be empirically studied via retrospective measures such as relative commercial success of samples of "first-to-market" and "second-to-market" functionally equivalent product pairs (the measure used by Sappho).

[16] "Lead time" is commonly defined as the period starting when an innovator introduces a new product to the market and ending when the first "me-too" product is introduced by a competitor. Lead time may be caused by any of the three innovation benefit capture mechanisms presented in this paper or by numerous other factors. Thus, an innovator may seek to prolong his lead time beyond the period afforded by response time by denying would-be imitators access to relevant know-how or patents and/or by various other means such as adopting pricing strategies designed to forestall imitation [35].

The value of response (and lead time) to would-be innovators can be reasoned to be a function of various situation-specific factors. One such factor is the length of response time divided by length of customer purchase decision cycle. A high value of this factor favors the innovator over imitators. Consider one extreme example: a consumer "fad" item (very short purchase decision time) which sells in high volume for six months only. Assume that the item can be readily imitated – but can only be produced economically by mass-production tooling requiring six months to build. Obviously, response time here allows the innovator to monopolize the entire market if he can supply it with his initial tooling. At another extreme is an expensive capital equipment innovation which customers typically take two years to decide to buy, budget for, etc. – and which competitors can imitate in one year. Obviously, response time in this instance affords an innovator little protection. A second situation-specific factor involves the learning curve: the more units produced during the response time period and the steeper the learning curve, the greater the production cost advantage an innovator can accrue relative to potential imitators. A third factor is the size and "indivisibility" of production plant investment an innovation requires relative to market size. For example, if DuPont uses response time to invest in a special-purpose plant for the production of Teflon which is large enough to supply any foreseeable market expansion for several years ahead, incentives to imitate are considerably reduced.

5. Differences in the ability of would-be innovators to appropriate benefit from innovation-related knowledge

In section 2 I concluded that it would be possible in theory to model the functional locus of innovation in terms of the appropriability of innovation benefit if would-be innovators: (1) *are not* able to capture benefit from non-embodied knowledge; (2) *are* able to capture benefit from output-embodied knowledge arising from their output-embodied innovations; and (3) differ significantly in their ability to capture benefit from their output-embodied innovation knowledge. In section 3 I concluded that the two mechanisms which an innovator might use to capture benefit from non-embodied knowledge (namely, patents and trade secret legislation) are relatively ineffective, that would-be innovators in most industries are therefore *not* able to capture benefit from non-embodied knowledge and that model condition 1 was thus satisfied. In section 4 I concluded that model condition 2 was satisfied on logical grounds: since would-be innovators can only appropriate economic benefit from their innovations by selling non-embodied knowledge and/or output-embodied knowledge, and since the former cannot be done effectively in most industries, most innovators *must* appropriate economic benefit from output-embodied knowledge if they appropriate such benefit at all. In this section I proceed to a consideration of the third condition which must be met if the functional locus of innovation is to be predicted in terms of the appropriability of innovation benefit – presence of a significant difference in the *ability* of would-be innovators having different functional relationships to a given category of innovation to capture benefit from output-embodied innovation knowledge.

Clearly, a difference in the ability of would-be innovators to capture benefit from output-embodied innovation knowledge must be substantial if one is to be able to predict the functional locus of innovation in terms of this single variable. I will begin this section by describing a study which "proves-by-example" – that the differences of the required magnitude can exist in the real world and then will offer some tentative generalizations regarding characteristics of industries likely to be associated with the presence of such major differences.

5.1. Differences in the ability to capture benefit from output-embodied innovation knowledge: the example

The proof-by-example that innovators holding different functional relationships to the same set of innovations can differ substantially in their ability to capture innovation benefit draws heavily on a 1977 study by Lionetta [38] of innovation in a plastics fabrication technology called "pultrusion". [17]

In the portion of the study of interest here, [18] Lionetta studied the machinery used in the pultrusion process from the invention of that process in the early 1950s to 1977. He identified all successful process machine innovations which offered the machine user "a major increment in functional utility at the time of its introduction when judged relative to best practice extant at that time". Lionetta next sought to determine the "source" of each such innovation via a careful search of contemporaneous literature and by interviews with user and manufacturer personnel who were involved in or found to have knowledge of the innovation work. In some instances he found the innovation involved the development of equipment unique to pultrusion. In other instances the innovation involved a first application of equipment used in other industrial processes to the pultrusion process. If we define the innovating firm as the firm which built the first unit of equipment embodying the (original or adopted) process machinery innovation which was used in commercial pultrusion production, we find that eight of the nine major process machinery improvement innovations samples identified were developed by machinery users and only one by a machinery manufacturer [40]. [19]

---

[17] Currently limited to the manufacture of fiber-reinforced products of constant cross-section, the pultrusion process is used to fabricate such products as the fiberglass-reinforced rod used by makers of fiberglass fishing rods. In essence, the pultrusion process involves pulling reinforcing material, usually fiberglass, simultaneously from a number of supply rolls into a tank containing a liquid thermoset resin such as polyester. The strands of reinforcement material emerge from the tank thoroughly wetted with resin and then pass through "preforming tooling" which aligns and compacts them into the desired cross-section. The compacted bundle of glass and liquid resin is then pulled through a heated die where the resin is cured and finally to a saw which cuts the continuously formed product into sections of the desired length. The entire pultrusion process is performed from start to finish on a single integrated machine. While the economic importance of this plastic fabrication process is still relatively small (only $60 million worth of "pultrusions" were produced in 1976), its use has grown at a real annual rate of 15–20% from 1967 to 1977, and some experts rank it second only to injection molding in terms of ultimate economic importance in the production of fiber-reinforced plastics [37].

[18] When a portion of Lionetta's study results were seen to be germane to the issues addressed in this paper, Lionetta was kind enough to join with the author in carefully cross-checking and updating the relevant subset. As a result, data presented here sometimes differ from the data presented in the 1977 study.

[19] This innovative machinery manufacturer, Goldsworthy Engineering, Inc., Torrence, Calif., was affiliated with a user of pultrusion machinery at Glastrusions, Inc., Torrence, Calif.,

In addition to determining the locus of process machinery innovation in pultrusion Lionetta examined the economics and structure of the US pultrusion machine user and pultrusion machine builder communities. He found approximately 40 firms using pultruders in 1976, producing an aggregate of \$60 million worth of pultruded product, at an average price of \$1.70 per pound and an average before tax profit of 12%. This product was produced on approximately 150 pultrusion machines,[20] each producing on the order of 200,000 pounds of pultrusions annually. Approximately 120 of these machines were found to have been "home-built" by the firms using them and only 30 to have been built by the only commercial builder, Goldsworthy Engineering, Inc. Pultrusion machine user firms were not able to supply useful data on the actual costs of the machines they had built over the years since they had often been built and rebuilt *ad hoc* by production engineers. However, Lionetta was able to estimate on the basis of data available on some recently built machines of "average" capacity (a machine capable of pultruding product with a cross-section of 6 by 7 inches) that a "home-built" machine of this capacity would have had a direct cost of \$50,000–60,000 in 1977, while company price lists show that an equivalent machine from the sole commercial builder would have had a purchase price of approximately \$95,000 at that time. Actual sales of commercially built pultruders were reported by the manufacturer to total four machines at an average price of \$35,000 in the years prior to 1967 and 26 machines at approximately \$105,000[21] each in the period from 1967 to 1977. Sales, therefore, of commercially produced pultrusion equipment in the 1967–77 period were on the order of \$270,000 per year. The manufacturer reported sales during this period to be relatively flat despite the annual real increase in annual output of pultruded product averaging 15–20%.

Lionetta's data can be used to construct a test of reason which strongly supports the proposition that one functional category of would-be innovators (in this instance, the users of process machinery in the pultrusion industry) have a much greater ability to appropriate benefit from output-embodied innovation knowledge related to their innovations than do those having other functional relationships to the innovation (in this instance, process machinery manufacturers).

Recalling the mechanisms for the capture of benefit from output-embodied innovation knowledge discussed previously, consider first the relative ability of pultrusion machinery manufacturers and users to capture benefit from such knowledge via the establishment of firm-level quasi-monopolies. The two mechanisms which we found likely to be effective in the establishment of such monopolies were response time and trade secrets (know-how). In the instance of pultrusion machinery process innovations it is clear that only user innovators can hope to retain control of their innovation related know-how much beyond the point at which commercial use begins. This is so because pultrusion process machinery innovations can be reverse engineered if inspected by would-be imitators skilled in the art. And, while an innovating machine user can exploit the innovation commercially while keeping it hidden from such inspection behind his factory walls, an innovating machine builder must make the innovative equipment available to the inspection of potential purchasers if he is to reap output-embodied benefit from it. This in turn opens the way to imitation delayed only by the response time of would-be imitators.

Two categories of trade secret are germane to the process machinery innovations being considered here – trade secrets bearing on the use of innovative equipment and trade secrets bearing on its manufacture. For reasons analogous to those spelled out in the previous paragraph, I conclude that only users are in a position to benefit from trade secrets regarding the use of innovative equipment, because only users can exploit these commercially while keeping them secret from would-be imitators. In contrast and again for analogous reasons, I conclude that both machine builder and

at the time of the innovation work. As part of a conservative stance toward the "user as innovator" hypothesis being tested, however, this firm's innovations were coded as machine builder developed.

[20] Lionetta obtained estimates from experts which ranged from 87 to 200 extant pultrusion machines. Through computations based on average machine capacities he developed an estimate of 176 machines.

[21] The difference between the \$95,000 list price for a commercial *equivalent* of an average home-built machine just noted in the text and Goldsworthy's average sales price is the inclusion of an optional RF curing unit costing on the order of \$25,000 in many of the units sold by Goldsworthy but not present on home-built machines.

machine user have a similar capacity to keep trade secrets regarding the manufacture of innovative equipment, and that both the single extant pultrusion equipment manufacturer and the larger users have similar incentive to develop such, as both build pultrusion equipment on approximately the same scale.

Consider next the relative ability of pultrusion machinery manufacturers and users to capture benefit from output-embodied innovation knowledge via the establishment of industry-level quasi-monopolies. Barriers to entry, the mechanism which allows the establishment of an industry-level quasi-monopoly, presumably provide some protection against new entrants to both machine builders and machine users in the pultrusion field. Lionetta's data show, however, that the machine builder apparently is unable to appropriate benefit from output-embodied innovation knowledge as a result of these barriers because as noted earlier, users, although they do not enter the commercial machine business, have proven themselves capable of building machines to satisfy their in-house needs at a cost at or below a machine builder's sales price for similar machines – presumably because the user does not incur selling expenses as the machine builder must. And in the pultrusion industry the machine manufacturer does not make significantly more machines than the largest users – and thus cannot offset these extra costs via economy of scale savings.

In contrast, it is very likely that machine users can appropriate benefit from output-embodied innovation knowledge via increased profits and/or sales as a consequence of an innovation-related industry-level quasi-monopoly. This is so because process machinery innovations in pultrusion typically allow the pultrusion industry to enter new markets at the expense of competing materials such as aluminum by making it possible to manufacture new shapes via this method. Thus, hollow product tooling, one of the innovations whose antecedents were examined by Lionetta, enabled pultrusion to be used to manufacture shapes of hollow as well as solid cross-section. Similarly, the development of improved "pulling" mechanisms, also examined by Lionetta, made it possible to pultrude shapes of larger cross-sectional area than had been possible previously.

Accordingly we may conclude that process machinery users in the pultrusion industry have a much greater ability than machinery manufacturers to appropriate benefit from output-embodied innovation knowledge derived from firm- and industry-level quasi-monopolies. One can illustrate this discrepancy quantitatively via Lionetta's data which show that an additional pultrusion machine employed by a machines user will allow the manufacture of 200,000 pounds of additional pultrusions annually. At the 1976 annual sales price of $1.70 per pound and pre-tax profit of 12% we can see that such an additional volume will yield the machine user $41,000 additional pre-tax profit *annually*.

In contrast, each extra machine sold by a machine builder as a result of the innovations is worth only a one-time profit of $10,000 to that firm at the prevailing machine price of about $100,000 and pre-tax profit rate of 10%. Thus, the machine builder would have to sell approximately four additional units annually as a direct consequence of his innovation in order to obtain benefit from his output-embodied innovation knowledge equal to that obtained by an innovating user who has embodied the innovation in only one machine and sold an extra 200,000 pounds of product thereby. Such an incremental volume on the part of the machine builder seems implausibly high given the sales rate of 2.6 machines annually which that firm has recorded during the 1967–77 period. On the other hand, embodiment of the innovation in only one machine seems an implausibly conservative estimate for the larger user firms since, as I have determined via telephone survey, the top three firms in the field had more than 15 pultrusion machines each in 1978.

In sum, then, I propose that condition 3 holds in this instance. That is, I have shown a strongly discrepant ability of firms holding different functional relationships (user, manufacturer) to the same class of innovations to capture benefit from output-embodied innovation knowledge regarding these. Further, data on the locus of innovation in pultrusion are in accordance with what I would predict if conditions 1, 2, and 3 of the single factor model are met.

5.2. *Toward generalization*

In the pultrusion industry I found that process machinery users and process machinery manufacturers had sharply discrepant abilities to appropriate benefit from output-embodied innovation knowledge regarding process machinery innovations. I found, further, that the cause of this difference could be logically attributed to mechanisms for the appropriation of benefit from output-embodied innovation knowledge identified and discussed earlier in this paper. Specifically, I found that the pultrusion process machinery manufacturer was not in a position to establish and benefit from an industry-level quasi-monopoly with respect to process machinery innovations because users could – and did – construct machines embodying the innovations at a cost competitive with the manufacturer's price when that price incorporated only a "normal" level of profit. In contrast, I reasoned that users might well establish an industry-level quasi-monopoly and that this mechanism of benefit capture was therefore either ineffective for both users and manufacturers or effective for users only. Next, I found that pultrusion process machinery innovations could be reverse engineered if inspected by persons "skilled in the art". Since only user-innovators are in a position to appropriate benefit from output-embodied innovation knowledge characterizing their innovations while secreting them from inspection within their factory walls, I concluded that user-innovators were more favorably positioned than manufacturer-innovators with respect to establishing and maintaining firm-level quasi-monopolies based on response time and trade secrets related to the use of a process machinery innovation. In contrast, both users and manufacturers were found equally favorably positioned with respect to establishing and maintaining firm-level quasi-monopolies based on trade secrets related to the construction of innovative process machinery.

Since most process machinery innovations can be reverse engineered if inspected by someone skilled in the art – and since most process machinery can be constructed on ordinary metalworking machinery available to would-be innovators in all functional categories alike – I propose that our pultrusion industry findings are generalizable to most process machinery innovations. That is, we may generally expect that all but one mechanism for the capture of benefit from output-embodied innovation knowledge will favor the user – because only users can appropriate such benefit from innovative process machinery while shielding the innovation from inspection by would-be imitators.[22] The sole mechanism *not* biased in favor of the user is, as was noted earlier, firm-level quasi-monopoly derived from trade secrets related to the construction of the process machinery innovation. Experience curve data indicate that the relative "amount" of this type of secret firms will acquire is a function of the relative number of machines they build.

On this basis one may venture the following economy-of-scale-related generalizations for all situations where users and manufacturers are the functional groups most favorably positioned to capture benefit from output-embodied innovation knowlege. When manufacturers of a given category of process machinery can reasonably expect to sell "many more" of a given process machinery innovation than any single large user can utilize, then process machinery manufacturers will be found to be the source of innovation in that category of process machines. Otherwise users will be found to develop – or pay for the development of – these. In a simple test of robustness of this generalization I interviewed process engineers at a razor blade manufacturer and a lamp manufacturer. In each instance the machine user firm was found to have developed and built the highly specialized equipment they required in-house. (An example of such equipment in the instance of the razor blade manufacturer was high-speed razor blade sharpening machinery and, in the instance of the lamp manufacturer, high-speed lamp assembly machinery). Both firms, however, were found to have purchased packaging machinery, used by many industries, from packaging machinery manufacturing firms.

The above generalization can be extended to explain why the locus of process machinery innovation might shift with time for some categories of process machinery, but not display such shift in others. Thus, the shift observed by Knight [42] from development of innovative computer hardware by users in the early days of that field to a later computer manufacturer locus of innovation is congruent with the single factor model. In contrast, I would not expect the locus of innovation to

[22] Note that there are exceptions to this user capability. For example, users of construction machinery used in the open clearly cannot shield innovations related to it "behind factory walls".

shift from user to manufacturer over time in the instance of razor blade sharpening machinery since the market for such specialized machines has been small in the past and will presumably remain so.

Although my own research to date on this variable has focused on the costs and benefits of certain categories of process machinery innovation, other categories of innovation look equally promising, and I would encourage investigation into many such. As noted earlier, Berger [5] and Boyden [6] have, for example, sampled plastics and plastics additive innovations respectively and have found all of these to have been developed by product manufacturers rather than product users. I suspect that further research would show this locus explicable in terms of the ability of users and manufacturers to appropriate benefit from output-embodied innovation knowledge in these categories of innovations. A particular plastic or additive is typically not essential to users since other materials exist which can do the job at a (usually minor) cost premium. To the manufacturer, however, a plastics and additive innovation which provides such a slight cost advantage may mean that major users of other materials (steel, aluminum, other plastics, etc.) replace these with the innovative material and quickly become major customers, thus allowing the innovator to capture significant benefit from output-embodied innovation knowledge.

In sum, I propose that the appropriability of benefit from output-embodied innovation knowledge is a variable which can usefully be incorporated in a model of the locus of innovation. I also propose that in some categories of innovation, not yet clearly delineated, the role of this variable in determining the locus of innovation is a strong one.

References

[1] M.J. Peck, Inventions in the Postwar American Aluminum Industry, in: The Rate and Direction of Inventive Activity: Economic and Social Factors, A Report of the National Bureau of Economic Research (Princeton University Press, Princeton, NJ, 1962), pp. 279-298.

[2] K.E. Knight, A Study of Technological Innovation: The Evolution of Digital Computers, unpublished Ph. D. Dissertation, Carnegie Institute of Technology, 1963. Data shown in table 1 of this paper obtained from Knight's Appendix B, parts 2 and 3.

[3] J. Enos, Petroleum Progress and Profits (M.I.T. Press, Cambridge, MA, 1962).

[4] C. Freeman, Chemical Process Plant: Innovations and the World Market, National Institute Economic Review.

[5] A. Berger, Factors Influencing the Locus of Innovation Activity Leading to Scientific Instrument and Plastics Innovation, unpublished S.M. Thesis, M.I.T. Sloan School of Management, June, 1975.

[6] J. Boyden, A Study of the Innovation Process in the Plastics Additives Industry, unpublished S.M. Thesis, M.I.T. Sloan School of Management, January, 1976.

[7] W.G. Lionetta, Jr., Sources of Innovation Within the Pultrusion Industry, unpublished S.M. Thesis, M.I.T. Sloan School of Management, June, 1977. When a portion of Lionetta's study results were seen to be germane to the issues addressed in this paper, Lionetta was kind enough to join with the author in carefully cross-checking and updating the relevant subset. As a result, data presented in the present paper sometimes differ from the data presented in the 1977 study.

[8] E. von Hippel, The Dominant Role of Users in the Scientific Instrument Innovation Process, Research Policy, 5 (1976) 212-239.

[9] E. von Hippel, The Dominant Role of the User in Semiconductor and Electronic Subassembly Process Innovation, IEEE Transactions on Engineering Mangement, May, 1977.

[10] K.J. Arrow, Economic Welfare and the Allocation of Resources of Invention in: The Rate and Direction of Inventive Activity: Economic and Social Factors, A Report of the National Bureau of Economic Research (Princeton University Press, Princeton, NJ, 1962), pp. 609-626.

[11] R. Nelson, The Simple Economics of Basic Scientific Research, Journal of Political Economy, 67 (June 1959) 297-306.

[12] A. Pakes and M. Schankerman, The Rate of Obsolescence of Knowledge, Research Gestation Lags, and the Private Rate of Return to Research Resources Discussion Paper Number 659, Harvard Institute of Economic Research, Harvard University, October, 1978.

[13] For a recent and extensive annotated bibliography of extant theoretical and empirical writings on this topic, see Institute for Defense Analyses, The Effects of Patent and Antitrust Laws, Regulations, and Practices on Innovation, 3 vols. National Technical Information Service (Arlington, VA, February, 1976).

[14] C.T. Taylor and Z.A. Silberston, The Economic Impact of the Patent System (Cambridge University Press, Cambridge, 1973).

[15] Ibid., p. 371.

[16] Ibid., see question 6 of section 6 of questionnaire, p. 396.

[17] Ibid., p. 30.

[18] Ibid., p. 196.

[19] F.M. Scherer et al., Patents and the Corporation, 2nd edition (Patents and The Corporation, J. Galvin, General Manager, Bedford, MA, 1959).

[20] The source of information in selecting the mail survey respondents was the list of corporate assignees in P.J. Federico, Distribution of Patents issued to Corporations

(1939-44), Study No. 3 of the Subcommittee on Patents, Trademarks, and Copyrights (Washington, DC, 1957), pp. 19-34.

[21] Scherer, op. cit., p. 107. A detailed breakdown of patent holdings of respondents is provided on pp. 108-109.

[22] Ibid., p. 117.

[23] Taylor and Silbertson, op cit., provides data in tables 8.4 (p. 153), 8.5 (p. 157), and 8.6 (p. 159) showing that only approximately half of the sampled license agreements in force in 1968 provided monetary return to the licensing firm in 1968. The distribution and amount of nonmonetary return possibly derived from these licenses in 1968 is not provided.

[24] R.W. Wilson, The Sale of Technology Through Licensing, unpublished Ph.D. Dissertation, Yale University, May, 1975.

[25] Ibid., Appendix, p. 152 ff., provides a detailed description of Wilson's data collection methods.

[26] Statement of Hon. David Ladd, Commissioner of Patents, before the Patents Trademarks and Copyrights Subcommittee of the Judiciary Committee, U.S. Senate, September 4, 1962, re: S.2225, as quoted in Elmer J. Gorn, Economic Value of Patents Practice and Invention Management (Reinhold, New York, 1964).

[27] See Carole Kitte, Patent Invalidity Studies: A Survey (National Science Foundation, Division of Policy Research and Analysis, January 1976), for references to and discussion of several such studies.

[28] D. Shapley, Electronics Industry Takes to Potting Its Products for Market, Science, 202 (November 24, 1978) 848-849. The practice of reverse engineering the IC designs of competitors is also discussed in How "Silicon Spies' get away with copying, Business Week, April 21 (1980) 178

[29] Those interested in examining the literature which does exist on trade secrets will find the following two articles a useful starting point: J. Mahon, Trade Secrets and Patents Compared, Journal of the Patent Office Society, 50:8 (August 1968) 536; L. Orenbuch, Trade Secrets and the Patent Laws, Journal of the Patent Office Society, 52 (October 1979) 659.

[30] See, for example, Enos, op. cit., for a detailed examination of the licensing of patents and know-how and the related revenue streams in one industry.

[31] M. Olson, The Logic of Collective Action: Public Goods and the Theory of Groups (Harvard University Press, Cambridge, MA, 1965).

[32] E. von Hippel, Transferring Process Equipment Innovations from User-Innovators to Equipment Manufacturing Firms, R&D Management, 8 (October 1977) 13-22.

[33] E. von Hippel, The Dominant Role of the User, op. cit., p. 14.

[34] J.E. Tilton, International Diffusion of Technology: The Case of Semiconductors (The Brookings Institution, Washington, DC, 1971).

[35] D. Gaskins, Jr., Dynamics Limit Pricing: Optimal Pricing Under Threat of Entry, Journal of Economic Theory, 3 (September 1971) 306-322.

[36] B. Achilladelis et al., Report on Project Sappho to SRC: A Study of Success and Failure in Industrial Innovation (Science Policy Research Unit, University of Sussex, Brighton, August 1971), 2 vols. Table 3.8, vol. 1, p. 92. (Instrument data on that table are incorrect; numbers have been transposed. See Appendix C. Complete Data Matrix, code 1, vol. 2, p. 58, correct data.)

[37] S.H. Pickens, Pultrusion - The Accent on the Long Pull, Plastics Engineering, July (1975).

[38] Lionetta, op. cit.

[39] Ibid., table 9, pp. 50-51.

[40] Ibid., table 10, p. 57.

[41] Ibid., p. 25.

[42] Knight, op. cit.