How do the Social Benefits and Costs of the Patent System Stack Up in Pharmaceuticals?

Daniel J. Gifford [*]

[*] Robins, Kaplan, Miller & Ciresi Professor of Law, University of Minnesota. The author acknowledges helpful comments on earlier drafts from Dan Burk, Dan Farber, Brett McDonald, David McGowan, Ruth Okediji, and Greg Polsky.

I. INTRODUCTION

This paper examines, from both theoretical and policy perspectives, a limited but important aspect of the patent system: its role and operation in supplying global demand for widely recognized health needs. It concludes that although the patent system is without peer in routing resources to the creation of the technological needs of modern societies, some aspects of that system operate better than others. In this connection, the paper directs attention to ways in which the patent system may produce less-than-optimum results in the markets served by the pharmaceutical industry as well as to related issues about how research on the world's widely recognized health needs should be funded.

The patent system, once largely ignored by non-specialists, has recently been receiving increasing attention from legal academics, economists, and policy makers. These analysts have focused both upon the system's domestic effects and upon its effects in the global economy. The creation, in the 1980s, of the Federal Court of Appeals with oversight over patent litigation,[1] brought renewed strength to the domestic patent system. Partly as a result of this reform, academic examinations of the system, which began in earnest in the 1960s, have increased dramatically. The negotiation of the World Trade Agreement in 1994 brought all of intellectual property into the world trading system through the ancillary TRIPS agreement,[2] subjecting it to new critiques from those sensitive to the impact of this property system upon the publics of the world's less developed regions.

Although some economists have been skeptical about the impact of the patent system in generating new technology,[3] others have recognized its potency. Perhaps Kenneth Arrow's 1961 inquiry into the differing innovation incentives found in concentrated and competitive markets[4] provided the initial spark for the substantial attention that the patent

---

[1] Pub. L. No. 97-164, 96 Stat. 36 (1982) (codified at 28 U.S.C. § 1295 (2000).

[2] Agreement on Trade-Related Aspects of Intellectual Property Rights, April 15, 1994, Legal Instruments of the Uruguay Round, 33 I.L.M. 81 (TRIPS Agreement).

[3] See C.T, TAYLOR & Z. A. SILBERSTON, THE ECONOMIC IMPACT OF THE PATENT SYSTEM (1973); Edwin Mansfield, *Patents and Innovation: An Empirical Study*, 32 MGMT. SCI. 173, 176 (1986).

[4] Kenneth Arrow, *Economic Welfare and the Allocation of Resources for Invention*, in THE RATE AND DIRECTION OF INVENTIVE ACTIVITY (1962).

system has received from economists. Later in that decade, William Nordhaus moved theoretical research a giant step forward with the publication of his seminal work on the economics of the patent system,[5] a work that stimulated an immense amount of analytical attention to the patent system and its operation. Edmund Kitch provoked the interest of legal scholars when, in the 1970s, he showed us how the patent system operates as a vehicle for staking out a particular area of technology for exclusive development,[6] a condition often critical to the investment of needed resources. Louis Kaplow drew the attention of the legal community to the costs and benefits of the patent system in his important 1984 work comparing the welfare effects of antitrust and patent market restraints.[7] Robert Merges took the lead in examining the operation of the patent system in a series of articles in the early 1990s.[8] Since Merges' pioneering work, legal scholars have joined others in a flood of works examining the patent system and its operation. Recently Mark Lemley and my colleague Dan Burk have provided a major contribution to this research with an examination of how the patent system operates in different industries.[9] Throughout this period, policy makers were generating new legislative modifications to the patent system. In the 1970s, congressional concern about the impact of time-consuming FDA review of new drug applications resulted in legislative extensions of the patent term for pharmaceutical companies that had lost initial years of patent protection to that review.[10] In the 1980s policy makers focused upon the patent system as an agent for economic rejuvenation, with the result that Congress created the Federal Court of Appeals.[11] And in the 1990s, Congress approved the NAFTA and WTO agreements that provided new strength to patents and other intellectual property rights

---

[5] WILLIAM NORDHAUS, INVENTION, GROWTH AND WELFARE, ch. 5 (1969).

[6] See Edmund W. Kitch, *The Nature and Function of the Patent System*, 20 J.L. & ECON. 265 (1977).

[7] Louis Kaplow, *The Patent-Antitrust Intersection: A Reappraisal*, 97 HARV. L. REV. 1813 (1984).

[8] Robert P. Merges, *Uncertainty and the Standard of Patentability*, 7 HIGH TECH. L.J. 1, 21 (1992); Robert Merges & Richard R. Nelson, *On the Complex Economics of Patent Scope*, 90 COLUM. L. REV. 839 (1990); Robert P. Merges, *Commercial Success and Patent Standards: Economic Perspectives on Innovation*, 76 CAL. L. REV. 803 (1988).

[9] Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, 89 VA. L. REV. 1575 (2003).

[10] Drug Price Competition and Patent Term Restoration Act of 1984, Pub. L. No. 98-417, 98 Stat. 1585 (1984) (codified as amended at 15 U.S.C. §§ 68b-68c, 70b, 21 U.S.C. §§ 301 note, 355, 360cc, 28 U.S.C. § 2201, 35 U.S.C. §§ 156, 271, 282 (2000).

[11] See note 1 *supra*.

throughout North America and the world.[12]

Currently the operation of the patent system is on the forefront of controversies, both domestic and international, about its effects upon pricing and exclusion in the pharmaceutical industry. We allow patentees to exercise exclusive rights—rights that may sometimes be equivalent to monopolies—over their inventions for a term of years precisely to create incentives to invent. And yet users of pharmaceuticals—especially the elderly—have complained so much about high pharmaceutical prices that Congress has legislatively reformed the Medicare Act to subsidize the purchase of pharmaceuticals.[13] The public policies that foster monopoly pricing in the patent law and those that subsidize purchasing in the amended Medicare Act appear to be in some tension. These Congressional actions are in further tension with the actions of Canadian and European regulatory schemes that are designed to place upward limits on pharmaceutical prices. They are in even greater tension with strongly held beliefs of third-world governments and their publics that the patent systems of the United States and other Western nations are depriving the world's poor of essential medications

This paper addresses the broad interplay between the incentive structure of the patent system and that system's social benefits and costs, viewed both on a national scale and, to a significant extent, on an international one. The paper examines the relation of private and social value to investment (and thus upon the basic economics of the system) with a view to identifying the system's weaknesses. It draws heavily from Louis Kaplow who developed a way of conceptualizing the marginal social costs and benefits of the patent system. It also draws from Kenneth Arrow who described incentives for innovation in competitive and monopoly contexts.

The paper compares the operation of the incentive structure of the patent system with other mechanisms for fostering inventive activity, as

---

[12] Uruguay Round Agreements Act, Pub. L. No. 103-465, 108 Stat. 4809 (1994); North American Free Trade Agreement Implementation Act, Pub. L. 103-182, 107 Stat. 2057 (1993).

[13] Pub. L. 108-173 (2003). Section 101 of the legislation, inter alia, adds subsidization of prescription-drug benefits to the Medicare program. See 18 U.S.C. § 1860D-2(b), as added by P.L. 108-173. The popular press is paying increasing attention to high drug prices in the United States vis-à-vis Canada and other developed nations. See, e.g., Roger Parloff, *The New Drug War*, FORTUNE, Mar. 8, 2004, at 144; *Why We Pay so Much for Drugs*, TIME, Feb. 2, 2004 at 4; Josh Benson, *Drugged*, NEW REPUBLIC, Nov. 7, 2003, at 12.

important background for its ultimate focus upon the relationship between the patent system and the generation of life-saving drugs. Although the paper readily concedes the general superiority of the patent system for eliciting inventive activity, the paper suggests that its superiority may not extend throughout the entire range of potential inventive activity. Indeed, the paper raises the question as to whether the patent system is superior in the context of pharmaceutical products that play (or could play) critical roles in the control of certain diseases or other disabilities.

The paper builds on the Kaplow perspective for assessing social costs and benefits. In so doing it attempts to articulate a perspective for carrying on the debate about the operation of the patent system and its application to pharmaceutical research. Drawing from that perspective, the paper raises at least two important policy issues especially connected with the marketing of pharmaceutical products and the fostering of pharmaceutical research. First, it raises the issue of price discrimination. Are laws, customs, or other practices discouraging or otherwise impeding the very price discrimination that could reduce deadweight loss and thereby increase social welfare? Second, when should public policy foster inventive activity through means other than the patent system?

Part II of the article reviews the standard incentive theory underlying the patent system. It summarizes the theory under which the patent law is said to harness the incentives of the inventor for the benefit of society. Part III examines the incentive structure, with particular emphasis upon two factors that affect the profitability of that research: (1) the probabilities that the firm undertaking the research will succeed in obtaining a patent for a commercially valuable result; and (2) the effects of the time lag between the period in which funds are committed to research and the period in which the results of that research produces revenue. Part IV employs the marginal analysis developed by Louis Kaplow to sketch out a schema for balancing social benefits against social costs, a schema that initially employs a linear analysis. Part IV also introduces the time dimension discussed in Part III into the analysis of social costs and benefits, concluding that as the patent term increases the rate of increase of social benefits slows while the rate of increase of social costs increases. Finally Part IV expands its schema by dropping the linear constraint from its model. With that modification, the model reveals that the proportionality between social benefit and cost that would accompany a linear model can be transformed, at least in theory, into a vastly

disproportionate relationship. Part V then raises the question of whether price discrimination can remedy these welfare problems. Part VI attacks the welfare problem from another angle. It inquires whether there may be a class of new pharmaceutical products for which financing schemes other than the patent system would better maximize aggregate welfare.

II. PATENT THEORY, MARKET FAILURE, ECONOMIC INCENTIVES AND SOCIAL WELFARE: THE BEGINNINGS OF ANALYSIS

A. In General.

The theory of patent law is straightforward. Society benefits from new technology. Yet in the absence of patent protection, invention would often go unrewarded. Unless a portion of this newly created economic value can be captured by its inventor, there is no incentive to innovate. Indeed, in certain cases there would be a negative incentive: invention often requires the expenditure of substantial resources in research and experimentation. This failure of the market to supply the incentive to invent is a result of a crucial absence of property rights.

When people provide goods and services, the property rights regime enables them to capture the economic value which they create by providing these goods and services. A producer or merchant owns the goods which are produced or provided. This enables him to trade the goods for compensation. In a similar way a service provider ensures that it provides services only on condition of being paid. The common-law property regime requires augmentation in those circumstances in which property rights do not provide a means for an inventor to capture at least a portion of the economic value which she has created. This void in the common law is filled by the patent law.

To an inventor who can meet its stringent standards, the patent law confers an exclusive right to make, use or sell the invention for a twenty-year period, commencing with the date on which the inventor files his patent application. Since the patent office normally takes one to three years in evaluating the patent and/or in negotiating with the patentee over the scope of his claims, the effective legal term may be more like seventeen years. For products like pharmaceuticals that require regulatory approval before marketing can begin, the effective period of protection

may be further reduced. The inventor's "reward", as it is sometimes called, consists in his exploitation of these exclusive rights. Her reward is thus determined directly by the receptivity of the market to her product. If the product is in high demand, then she is likely to profit handsomely. Yet however ingenious the invention, there is little or no reward to the inventor unless buyers appreciate it and are willing to pay for it.

This dependence upon the combination of the exclusive rights conferred by the patent law with the incentives of the market has both positive and negative effects. On the positive side, the system ensures that incentives are directed towards the generation of products that people want. On the negative side, the patent system does not provide incentives for products for which there is a social need but no economic demand, such as drugs for diseases (like sleeping sickness) that primarily affect populations with little purchasing power. The patent system, almost by definition, also does not work to stimulate primary research. In these latter areas (needs of the poor, primary research) alternative systems of stimulating research and/or invention, such as by government funding[14] or by post-hoc government rewards are, or may be, necessary.[15] The patent system also generates inefficiencies: the patentee's exclusive rights permit it to charge supra-competitive prices for the patented product, with the result that some potential customers who value the invention at more than its cost of production but at less than the price charged by the patentee go unserved. In the language of economists, this is a deadweight loss: a loss to society resulting from a misallocation of resources.

Because the patent system operates through harnessing market-based incentives, the structure of those incentives bears examination. It is the expected value of the patented invention that provides the incentive to undertake the research and development that ultimately produces it. Before a potential inventor commits an investment to research and development activities, it assesses the expected profit from that investment. And, of course, it compares that expected profit with expected

---

[14] Most basic research is funded by the federal government. SEE STATISTICAL ABSTRACT OF THE UNITED STATES: 2001 508 (Table no. 769)

[15] Numerous suggestions have been made for a reward system of stimulating research. The focus of many of these proposals has been the elimination of the dead-weight loss problem. See, e.g., Michael Abramowicz, *Perfecting Patent Prizes*, 56 VAND. L. REV. 115, 122-27, 169 (2003); Steven Shavell & Tanguy Van Ypersele, *Rewards Versus Intellectual Property Rights*, 44 J. LAW & ECON. 525, 529 (2001).

returns from alternative investments.

B. The Patent System as an Adjunct to the Market

Although the patent system is not the only means available for fostering invention, it possesses certain characteristics that enable it to mesh with the market more or less seamlessly. We observed earlier that the patent system solves a market failure: the absence, in the traditional property regime, of rights over inventions means that the economic incentives that elsewhere foster productive behavior would not (in the absence of the patent system) foster inventive behavior. By providing these missing rights, the patent system broadens the reach of the market, endowing it with a major responsibility for stimulating invention in both end products and technology.

It is in this augmentation of market mechanisms that the advantages and disadvantages of the patent system lie. By providing missing property rights and relying upon the market to provide both the inventive stimulus and ultimate reward, the patent system maximizes the extent to which new inventive activity will be directed to the needs and wants that society most values (as measured by market demand) and minimizes the extent to which resources will be directed to unwanted goods and services. Throughout the operation of the patent system, the market plays the key role. Potential inventors look to the market for clues as to what kinds of products are likely to be rewarded. They gear their efforts according to the clues that the market provides. And the extent to which they are in fact rewarded for their inventive activity is determined by the market. The objective forces of the market thus perform critical roles in directing the course of inventive activity. Because no other decision-making mechanism can match the market's predictive abilities or its ability to continually reassess and reevaluate, the patent system, which incorporates these market mechanisms, partakes of these advantages. The superiority of the patent system over alternative means of fostering inventive activity thus lies in its ability to harness the powerful forces of the market to its ends. Yet it is also this attachment of the patent system to market mechanisms that account for its disadvantages.

C. Classes of Inventions and the Loci of the Patent System's Advantages and Disadvantages.

In a well-known paper of 1962, Kenneth Arrow divided inventions into two classes.[16] In the first class are innovations which reduce production costs substantially. In the second class are innovations which reduce costs in lesser amounts. The first class embraces innovations which lower cost so much that—were the market to be in control of a monopolist—that monopolist would set the post-innovation profit-maximizing price below the level of the old unit production cost. The second class consists of other cost-reducing innovations. This classification of inventions worked well for Arrow's paper which sought to distinguish the profit generated by invention in a monopoly marketplace from that generated by invention in a competitive marketplace, and his classification has has been followed by others. Nordhaus, for example, employed that classification, and called the first class "drastic" inventions. Arrow's classification also works for this paper. Drawing from (and somewhat modifying) traditional legal terminology, this paper calls these two classes of inventions "pioneer" inventions[17] and "improvement" inventions.

The patent system probably operates in its least controversial mode in fostering improvement-type inventions. Here the ratio of dead-weight loss to profit is minimized. Minimizing this ratio mutes controversy over the optimum length of the patent term in the context of improvement-type inventions. And the fostering of improvement-type inventions shows the operation of the patent system at its best. The benefits of the system's decentralized incentive structure ensure that adequate attention is directed to improvements of technologies at levels that fall below the threshold of public visibility but which in the aggregate contribute significantly to the improvement of society's productive efficiencies. Probably most patent activity is concerned with improvement-type inventions. If so, then most patent activity is concentrated where it raises few controversial issues about social costs and benefits.

---

[16]Kenneth Arrow, *Economic Welfare and the Allocation of Resources for Invention*, in THE RATE AND DIRECTION OF INVENTIVE ACTIVITY (1962).

[17] Cf. Westinghouse v. Boyden Power Brake Co., 170 U.S. 537, 561-62 (1898) ("This word [pioneer invention] . . . is commonly understood to denote a patent covering a function never before performed, a wholly novel device, or one of such novelty and importance as to mark a distinct step in the progress of the art, as distinguished from a mere improvement or perfection of what had gone before.") In using the term "pioneer invention" to refer to a major invention in the Arrow sense, the text is employing that term in a related, but slightly different, sense from that used by the courts.

The patent system's most apparent disadvantages involve the deadweight losses that the system generates by conferring market power on patentees. These losses may be a part of a system that generates inventive activity, but they are nonetheless a social cost. Pioneer inventions are likely to generate higher ratios of deadweight loss to profit than improvement-type inventions. In some cases, the ratio of deadweight loss to profit might be very high. As a result pioneer inventions better raise issues about the system's social costs and benefits. Of course, the deadweight loss generated by pioneer inventions is also a measure of the social value created by these inventions. Society wants and needs pioneer inventions. The questions are whether patent terms are too long and whether these inventions can be generated with a lesser degree of deadweight loss.

The public is probably most conscious of patentee market power over new pharmaceutical products. Because pharmaceuticals are one of the few places where a single patent covers an entire product,[18] they may be less subject to pricing constraints than other inventions that are improvements to machines or processes and for which pre-existing technologies are ready substitutes. The media has reported extensive public concern over what are perceived as unduly high price levels for patented pharmaceuticals, a concern to which Congress has recently responded. Beyond these domestic welfare and distributional issues, however, the pricing of patented pharmaceutical products appears to create extensive deadweight losses in third-world nations. These real and perceived disadvantages of the patent system, as it operates in the pharmaceutical industry, may be accompanied by some weakening of the informational advantages that the system draws from its close interaction with the market. The system's ability to harness market-supplied information that is one of its major advantages may be less so with certain kinds of pharmaceutical products, where in the area of critical and life-

---

[18] See Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, 89 VA. L. REV. 1575, 1590 (2003):

> In some industries, such as chemistry and pharmaceuticals, a single patent normally covers a single product. Much conventional wisdom in the patent system is built on the unstated assumption of such a one-to-one correspondence. . . . Such a correspondence is the exception rather than the rule, however. Machines of even moderate complexity are composed of many different pieces, and each of these components can itself be the subject of one or more patents.

See also Herbert Hovenkamp, Mark Janis & Mark A. Lemley, *Anticompetitive Settlement of Intellectual Property Disputes*, 87 MINN. L. REV. 1719, 1738 (2003).

saving drugs, needs are widely recognized. Indeed, the patent system's close interaction with the market explains why pharmaceutical companies do not develop drugs for the cure of diseases afflicting poor nations. Such products would not be profitable. None of these remarks is meant to say that the patent system does not work in this industry or that it works particularly badly. Neither is it to say that social costs outweigh benefits in the generation of new pharmaceuticals. Rather, it is merely to point out that in this area the patent system's advantages appear less strong than they do in other areas. These apparent weaknesses in the way the system operates in pharmaceutical markets are discussed below. The article then discusses means for minimizing those weaknesses.

III. THE INCENTIVE EFFECTS FROM THE PERSPECTIVE OF THE INVENTOR: WHERE SOCIAL BENEFIT MEETS PRIVATE INCENTIVE[19]

A firm contemplating research to develop a new product necessarily investigates whether the research is likely to succeed and whether the revenues that the product generates will be likely to cover its costs and to product a profit. The ordinary lag between the time when a firm introduces a new product and the time when its competitors bring rival products to market provides a window for the innovating firm to capture much of the economic value that it has created. That period of de facto exclusivity is sufficient to support modest research and development. The patent system provides legally protected exclusivity for the longer periods necessary to justify the larger investments that may be necessary to design highly innovative products.

Thus a firm contemplating a large research investment considers, first, the chances that its research will succeed. Second, it considers the probabilities that it (rather than one of its rivals that may also be conducting research) will be able to obtain a patent on the product. Third, it assesses the amount of expected revenue that the product is likely to generate and the costs that it will incur in producing the product.

---

[19] The discussion in this and the next Part considers the value of the patented invention. We distinguish between the aggregate social value of the invention and the private value to the inventor. In the next Part, we consider the social value of invention. Here, because we are interested in the incentive effects of the patent system, we are concerned with the private value to the inventor.

A. The Probability of Success

A potential innovator must, of course, balance the amount of its research and development costs and its probability of successfully developing the innovation against the value of the innovation. In addition, it must also weigh the risk that a rival will also succeed in developing the invention. Professor Robert Merges breaks the decision about committing funds to research into two stages. In the first stage, the inventor decides whether to undertake preliminary experimentation on an invention. In the second stage, the inventor decides whether to develop the invention.[20] As Merges points out, this model captures some of the complexity of the real world: The results of the preliminary experimentation in the first stages provide information that will recast the probabilities of success that the inventor weighs when deciding whether to proceed into the second stage. Indeed, an inventor is continually facing new decisions with increasing amounts of information as the project proceeds. Thus, Merges' analysis fits nicely with a third stage of the development process identified by Edmund Kitch.[21] In Kitch's model, the issuance of a patent is treated as tantamount to staking out an area for commercial development. In this post-patent stage, the inventor has solved the basic technology problem and has won the race to the patent office. At this late stage, an inventor deciding whether to go forward must weigh the costs of commercial development against his estimate of commercial success.[22]

Let's consider the probability of success in the quest to develop the innovation. If only one firm is in the research race, then that firm undertakes the research if the value of the anticipated product times the probability of successfully developing it is greater than the cost of the required investment in r&d.[23] But if two firms decide to undertake

---

[20] Robert P. Merges, *Uncertainty and the Standard of Patentability*, 7 HIGH TECH. L.J. 1, 21 (1992).

[21] See Edmund W. Kitch, *The Nature and Function of the Patent System*, 20 J.L. & ECON. 265 (1977).

[22] The present discussion collapses the two stages identified by Merges for ease in presentation. Moreover, it also combines inventor's assessments about the probabilities of succeeding in his quest for a patented invention with his assessments of succeeding in the race to the patent office. Combining these several issues confronting an inventor here is not problematic so long as we remember that we are employing a simplifying model of a process that in real life involves a series of decisions, each of which draws upon a body of information that is continually being augmented.

[23] The analysis employed in this and the following footnotes is drawn from OZ SHY, INDUSTRIAL ORGANIZATION 224-25 (1995). Let: V = the private value of the invention; I = the anticipated cost of research and development; ; α [$\alpha < 1$] = the probability of technical

investment in similar r&d, the dimensions of the problem change. If one firm develops the product and the second firm does not, the successful firm acquires the entire value of the product.[24] But it is possible that both firms may succeed technologically but only one of them wins the race to the patent office and thereby captures the potential economic value of the product. Indeed, regardless of how many firms succeed technologically, only one will receive the patent.[25]

The expected value of product development must take into account both situations. Since only one firm can win the patent race, the probability of succeeding overall, i.e., the probability of both developing the technology and receiving the patent, is the multiple of the two probabilities. (i.e. probability of developing the technology x probability of receiving the patent). Accordingly, the ex ante expected profit for each firm is discounted by both probabilities. Of course, any number of firms may enter the research-and-innovation race. If we assume that each of the successful innovators has an equal chance to obtain a patent, then the expected value of the patent becomes the value of the innovation divided by the number of successful innovators. As the number of firms focusing

---

success; 1 - α = the probability of failure. If only one firm seeks to develop the product, then it will undertake the investment if its expected profit (π) is αV > I. Note that V is less than the social value of the invention, which includes consumer as well as producer surplus. V here is the capitalized value of producer surplus.

[24] The probability that one firm develops the product and the second firm does not is: α(1 - α).

[25] It is necessary to modify Shy's analysis somewhat when we deal with the probability that both firms succeed technologically, but only one firm succeeds in obtaining the patent. In the two-firm case discussed by Shy, the two share the product's value when both succeed. See *id*., at 224, Assumption 9.1. This, of course, cannot be literally true. If the product is patentable, the first inventor captures the entire value of the invention. Shy's treatment is acceptable for analytical purposes, however, because treating the rival firms as sharing the value of the invention is tantamount to according each of the rival firms an equal probability of winning the race to the patent office. The probability of this result is: α^2. The expected value of product development must take into account both situations. Accordingly, the expected profit for each firm is: $\pi = \alpha(1 - \alpha)V + \alpha^2V/2$. Each firm will undertake the requisite r&d if $\alpha(1 - \alpha)V + \alpha^2V/2 > I$, or equivalently if $\alpha(2 - \alpha)V/2 > I$.[25] Again, analysis shows that in areas of high cost r&d, firms will be increasingly reluctant to undertake r&d as the number of equally competent rivals in the research race grows. With three firms the potential payoff is: $\pi = \alpha(1 - \alpha)^2V + 2\alpha^2(1 - \alpha)V/2 + \alpha^3V/3$. Each firm will undertake the requisite r&d if: $\alpha(1 - \alpha)^2V + 2\alpha^2(1 - \alpha)V/2 + \alpha^3V/3 > I$. In its more generalized form, the expected profit for each firm doing research for product development grows in complexity:

$$\pi = \sum_{I=1}^{n} \frac{(n-1)!}{(I-1)!\,(n-I)!}\, a^{i}\,(1-\alpha)^{n-I}\,V/I$$

their efforts on developing the same innovation increases, the lower is the chance of success for any one firm. Indeed, as the number of firms in the patent race increases, the expected value of the invention to any one of them approaches zero.[26]

B. Returns Discounted to Present Value.

The expected profit that is salient to the inventor is, of course, the discounted present value of the expected future returns. As a result, the more distant are the future revenues, the lower is their contribution to the incentive structure of the patent system.

1. Incentives and Discounted Future Revenues.

Because the expected profit is realized over a period that begins only after the invention is fit for commercial exploitation, the comparison of expected profit to investment—as noted above—necessarily is a comparison of present costs with future earnings. Earnings necessarily are weighted less than the research costs that generate them because the former must be discounted to present value while the latter do not.

Moreover, the anticipated returns must be adjusted in several ways. We have already noted that they must be adjusted for the estimated probabilities of technological success as well as for the probabilities that technological success may be rendered moot by others winning the race to the patent office.

In this section, we focus upon the value of the anticipated returns. At the stage at which the investment commitment is made, the present value of the anticipated returns is at its lowest. Since the research and development work has not yet begun, the period in which the anticipated returns are generated is still some time away. There can be no returns until the invention is produced and in form for commercial use. Thus if the research and development period takes, say, three years, then the first returns will not begin until then. Thus, in that case, the first year's anticipated returns must be discounted for that three year wait. And, of course, the returns generated in each year of the patent term have to be discounted accordingly. (If the patent application is filed three years after

[26] See note 25 *supra*.

the commencement of research and the product is immediately marketed, the returns to the investor (viewed from the date of his investment) from each year of the 20 year patent term would have to be discounted from 3 to 23 years).

Kitch's focus upon patent rights as means of staking out a technology for commercial development[27] calls our attention to the fact that the early years of the patent term may not be usable for commercial exploitation. However long this post-patent period of development extends, it consumes some of the protected patent period, narrowing further the period of return on investment.[28] Modifying the above example to take account of a period of development, we might have a three-year period of research followed by filing of a patent application, which is then followed, say, by a two-year period of development. Since the 20-year patent term begins with the filing of the patent application, the two-year development period in the example reduces the commercially-relevant protected term to 18 years. So, in such a case, the first return would be a full five years distant from the commencement of investment. This problem became acute in the pharmaceutical industry after Congress required new drugs to be "effective" as well as "safe." The new effectiveness requirement added to the delay before a patented drug could be marketed, as more extensive testing was needed before the FDA could approve its marketing. Congress then authorized extensions of the patent term to compensate for regulatory delay, but the extensions do not fully compensate for those delays.

Finally, it should be observed that it may take time for potential customers to recognize the value of a new product and to adjust their purchases accordingly. A producer generally plans a promotional strategy with which to acquaint potential purchasers about the characteristics of a new product. As a result the sales volume of a new product may increase over a period of years. Revenue, accordingly, may be lower in the early years in which it is marketed than in later ones.

---

[27] See Edmund W. Kitch, *supra* note 21.

[28] As observed above, Merges' model identifies a first stage involving preliminary experimentation, followed by a second stage in which the inventive work is pursued. Merges, *supra* note 20. In this second stage, the risks are whether the technology can be developed at all and whether a rival will do it first. The goal of the latter stage is patentability, independent of commercial viability. In Kitch's third stage, the risk is solely with whether the technology can be raised to a level of commercial viability. The threat from rivals has been avoided by the patent obtained at the end of stage one.

2. The Inventor's Perspective Again.

Implicit in the discussion above are the economics underlying the decision of the inventor about whether to undertake an investment in innovation. The basic economic questions are whether the value of the patented invention is expected (1) to exceed the cost of the investment and (2) to produce a return superior to alternative investments. In making these comparisons, the prospective investor necessarily compares the present cost (the up-front investment) with future returns, returns that must be discounted both for their uncertainty and their future dates.

Let's illustrate these matters using a stylized example. Suppose we estimate the chances of successfully developing a new product (let's call it a widget) at 80%. Then we should discount our projected profits by 80%. Suppose further that we know that three of our rivals are attempting to develop this product. The first to succeed will receive a patent and block the others from the widget market. Since we and our three rivals are starting out about the same time and with approximately the same resources, our chances of developing the product first would appear to be one in four, or 25%. Thus, on this assessment, we have a 25% of 80% chance of success, or an overall probability of success of 20%. On these probabilities, unless the expected return is extremely high and we are high-risk takers, we should probably look for an alternative line of research and development.

Let's add a new element. Our own prior research gives us an advantage unknown to the others. As a result, we have concluded that we have a 90% chance of developing the product first. Now our aggregate probability of success is 90% of 80%, or 72%. The project, of course, is risky, but if the expected returns are sufficiently high, then they can justify the risk. Those returns must be discounted to 72% of their expected value in order to compare them with our investment and alternative investments. (Alternative investments, of course, also have to be discounted for risk).

Let's make some additional simple assumptions assume for illustrative purposes. We expect that our invention will generate revenues (in excess of production costs) of $1,000,000 per year. Over a twenty-year patent term, the total return would be $20,000,000. As we have noted, however, the ex ante expected value of the patented invention is substantially less than $20,000,000. Ignoring, for the moment, the time

required for research and development, that $20,000,000 in future earnings would have to be discounted to present value. Let's assume that we recognize earnings at the end of each of the twenty years of the patent period. Then the $1,000,000 for the first year is discounted by the interest rate. [$1,000,000/(1+r)], where r is the rate of interest.] The $1,000,000 for the second is discounted by the interest for years one and two. [$1,000,000/(1+r)^2]. And so on. The present value of the $20,000,000 in future earnings thus is:

$$\sum_{i=1}^{20} \$1{,}000{,}000/(1+r)^i$$

At an interest rate of 5%, the $20,000,000 in future revenues would have a present value of $13,850,320. So the present value of these revenues is only about 69% of their nominal dollar amount. We also must discount the $13,850,320 for risk. Recalling that we had an estimated 72% probability of succeeding in actually acquiring the invention, the value of the expected invention is $9,972,230. That is not quite 50% of the total expected future $20,000,000, expressed in nominal dollars. If the investment required to develop the product is substantially less than the $9,972,230, then it would provide a positive profit.

Suppose the required investment is $5,000,000. Then, on these figures, the investment would generate a profit of $4,972,230 over a twenty year period. The attractiveness of that return depends upon its alternatives. Five million dollars invested at 6% over a twenty year period would produce $19,098,748, or a net profit of $14,098,748 in nominal dollars or $5,054,202 in present value. Thus in this case, the pursuit of the invention does not appear especially attractive as an investment. An alternative disposition of the $5,000,000 at 1% over the going interest rate would be more profitable.

Thus the incentive structure of the patent system requires that we assess the expected return in the light of the investment necessary to generate that return. To make that assessment, future revenues must be discounted to present value and be further discounted for risk. Such discounting is standard practice for investors, considering whether or not to undertake a particular investment or in selecting a particular investment from a range of alternatives. Yet the patent system is dedicated to generating significant advances: No patent can be issued unless the

invention is "nonobvious,"[29] a phrase meaning beyond the knowledge and abilities of a competent professional in the field. The system itself thus courts the risk that the purported inventor will not be able to exceed the capabilities of his or her peers. In the most successful inventions, the risk factor will be reflected in the high profits which those inventions command. And the reliance of the patent system upon incentives generated by twenty-year period of exclusive rights, also exacerbates the difference between the return seen by observers (the dollar return at the moment of the observation) and the incentive to the inventor (that return discounted to its ex ante value). As we will see below, these differences between the ex post and ex ante values are relevant in a variety of ways to the assessment of the system's private and social costs and benefits.

IV. EXPLORING SOCIAL COSTS AND BENEFITS

A. The Arrow Analysis: Different Values in Different Market Structures.

Arrow (who divided inventions into two categories[30]) concluded that for pioneer inventions an innovator would set a royalty equal to a monopoly return. This would be true regardless of the industry market structure. Despite the monopoly return to the inventor, the public would incur an immediate benefit because the large cost savings would press prices downward, below the level that they were prior to the invention. This easily seen in figure 1 below. Let's assume that the product was initially produced in a competitive market where the cost and price are represented by the line PP and the output is XX. The invention reduces costs to the level of c. Now the inventor licenses the invention at a per unit royalty equal to the vertical distance between c and p. Consumer surplus is the area under the demand curve down to P, i.e., it is the area within the triangle DAP, albeit the contribution to consumer surplus of the invention is not the entire area of the triangle DAP but the lesser area of the trapezoid PP-Z-A-P. Similarly, the social benefit produced by the invention is not the area of the triangle DCG, but the trapezoid PP-Z-A-B-C. The inefficiency generated by the patent-system is represented by the triangle AGB, representing deadweight loss, i.e., the demands of those unsatisfied customers who value the product more than its production

---

[29] The patent act requires that an invention pass the threshold of obviousness. That is, the difference between the invention and the prior art must not be obvious to a person skilled in the art pertaining to the subject matter of the patent. 35 U.S.C. § 103 (2000).

[30] See *supra* note 16 and accompanying text.

cost (0c) but less than the patentee's price (0p).

The second or "improvement" category requires some discussion. Arrow was interested in how market structure affected the incentives to innovate. Although his focus differs from that of the present paper, his analysis is useful for exploring the economic effects of an invention of the improvement type. Arrow observed that in the case of an improvement, the inventor would set the royalty in the full amount of the per-unit cost savings, thereby capturing the entire cost savings for itself. But the revenues earned would differ depending upon the market structure in which the invention was employed. In a competitive market, the royalty would equal the unit cost savings multiplied by the industry output immediately prior to the deployment of the invention. In a monopoly marketplace, the inventor's return would be less, because the value of the invention would be limited to the unit cost savings multiplied by the smaller monopoly output. The critical element in Arrow's analysis of improvement patents, for our purposes, is the constraint that the preexisting technology exerts upon the patentee's pricing power. We examine that constraint below.

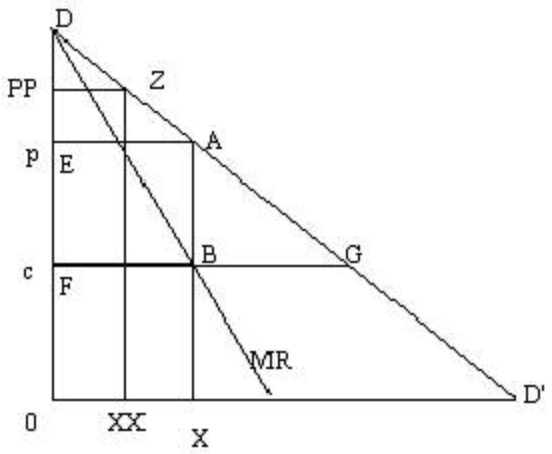

Figure 1

Figure 2 below depicts the Arrow hypothesis in diagrammatic form. In the diagram, the monopolist's initial profit is represented by the rectangle formed by the horizontal line p_{m1} on the top, the horizontal line c_1 on the bottom, the (vertical) 0 axis on the left side, and the vertical x_{m1} intersect on the right side (or $x_1(p_{m1} - c_1)$). The monopolist's post-innovation profit is represented by the rectangle formed by the horizontal line p_{m2} on the top, the horizontal line c_2 on the bottom, the (vertical) 0

axis on the left and the vertical x_{m2} intersect on the right (or $x_{m2}(p_{m2} - c_2)$). The increment to the monopolist's profits resulting from the innovation, accordingly, is the difference between these two amounts. Another way the difference in monopoly profits is shown in the diagram is in the area under the marginal-revenue curve (MR). The pre-innovation profits are represented by the area under the MR curve down to the initial cost curve c_1. The post-innovation profits are represented by the area under the MR curve down to the post-innovation cost curve c_2. The increment to the monopolist's profits from the innovation thus are represented in the diagram by the area under the MR curve between the two cost curves c_1 and c_2. This way of representing the increment to the monopolist's profits makes it easy to show Arrow's point graphically. When the monopolist innovates, it can capture additional profits represented by the area between c_1 and c_2 that lies under the MR curve. But when a competitor firm innovates, it can capture additional profits represented by the area between c_1 and c_2 all the way out to the pre-innovation competitive output of x_{c1}.[31]

The preexisting technology—represented here by c_1—constrains the royalty that the patentee is able to charge. Improvements can range in significance all the way from one that approaches (but does not reach) the cost-savings of a pioneer invention to the probably more common inventions that generate modest cost savings. In the context in which we have so far directed our attention, that of linear demand and constant costs, all improvement-type inventions generate a lower ratio of deadweight loss to profit (and to total surplus) than do pioneer inventions. Restated, within the context of linear demand and constant cost, this class of inventions appears prima facie to generate a higher ratio of social benefit to social cost than does the class of pioneer inventions.

---

[31]As pointed out above, the competitive output remains at x_{c1} because the innovator imposes a royalty for the newly developed technology equal to the unit cost savings. Without the royalty, the competitive output would increase to x_{c2}.

Arrow's conclusion that a competitor firm possesses a greater incentive to innovate than a monopolist rests upon an analysis that does not take account of how the incentives of a competitor to innovate are affected by the presence of rivals who are also considering the pros and cons of r&d activity. A secure incumbent monopolist (in the Arrow sense) need not discount its likelihood of technical success by the probabilities of beating its rivals to the patent office as must a competitor firm.

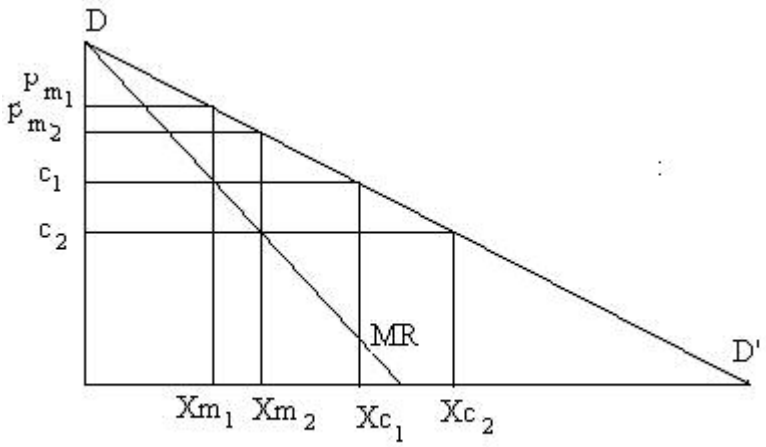

Figure 2

B. The Kaplow Analysis.

In a path-breaking analysis of the patent system and its connections to antitrust law, Harvard Professor Louis Kaplow brought a new analytical refinement to the evaluation of the patent system and its operation.[32] Identifying the social benefits of the patent system as the innovations that it engenders and the social costs of that system as the monopoly output restrictions which the system provides as incentives for innovative activity, Kaplow directed his attention to the system's marginal benefits and costs. As Kaplow rightly indicated, a rational society--after determining the level of innovation that it desired--would wish to generate that innovation at the least cost. Indeed, ideally society should limit the patent term at the point when the marginal social costs imposed by the patent system rise to the level of the marginal benefits that it generates.

In trying to get a handle on the patent system's marginal social costs and benefits, Kaplow hypothesized a one-year extension of the patent term. A one-year extension of the patent term would increase the reward of the patent and thus would probably generate more innovations. The additional innovations generated by that one-year extension would constitute the marginal benefit of such an extension. But a one-year extension of the patent term would also impose additional social costs. That one-year extension would mean that all of the patent monopolies which were about to expire now would continue for an additional year.

---

[32]Louis Kaplow, *The Patent-Antitrust Intersection: A Reappraisal*, 97 HARV. L. REV. 1813 (1984).

That additional year extension thus would impose monopoly losses upon society which society would not incur in the absence of the extension. The monopoly loss so imposed during the one-year extension of the patent term would be the incremental, i.e., marginal, cost of that one year extension. Kaplow also worked backward to compare the costs and benefits which would result from reducing the patent term: what would be the social losses from the reduction in innovation which would result from a one-year reduction of the patent term? And what would be the social benefits (in the elimination of monopoly restrictions) which would result from such a one-year reduction of the patent term?

Kaplow forthrightly acknowledged that it is virtually impossible to determine either the value of new innovation or the monopoly loss from a hypothetical extension of the patent term. He believed, however, that the analytical format which he developed would be helpful in thinking about the issues. His marginal analysis is a major contribution. Earlier writers had not focused their attention on marginal costs and benefits. Kaplow's format, by forcing us to consider marginal costs and benefits, moves analysis and evaluation of the operation of the patent system to a higher plane of conceptual clarity.

C. A Simple (and Linear) Model

Let us now use this marginal analysis in assessing the operation of the patent system. As Kaplow pointed out, when we protect an inventor's work for more years than would be necessary to stimulate his invention, we impose a cost upon society in the form of a restricted output. Thus some inventions upon which patents are granted would have been produced, even without a stimulus from the patent system. For other inventions, the costs of research and development, the risks of failure and the risks of the marketplace would deter the necessary innovative effort without the stimulus of an exclusivity period provided by the patent system. Yet for some inventions, the necessary period of exclusivity might be very short. The stimulus necessary to generate other inventions might be longer, but still less than the actual patent term. For yet other inventions, exclusivity for the full patent term would be necessary to provide an adequate stimulus. Other potential inventions may well go uninvented because even the 20-year patent term is too short to provide the requisite incentive.

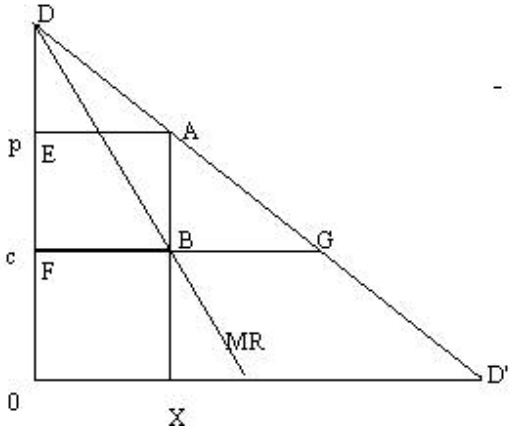

Figure 3

Let's begin our analysis by assuming that some inventions would be produced without any period of exclusivity, some would be produced with a one year period, others with a two-year period, others with a three year period, etc. Thus we will assume that each year of the patent term produces an incremental stimulus to generate more inventions. More precisely, we assume that invention occurs in proportion to the investment in research and development (r&d). On these assumptions, more investment in r&d generates increased invention and the increase in invention is proportional to the increase in r&d investment. For purposes of the following discussion, invention is measured in terms of a value derived from the market: Inventions carry value determined by the demand for a patented invention, a value derived from its attractiveness as a consumption good or from its ability to produce new products or to lower production costs for preexisting products.

Let us assume that each year of the patent term provides the incremental incentive necessary to generate new inventions, over and above the inventions generated by terms of lesser length. For purposes of exposition, we hypothesize that each year of the patent term generates such additional inventions that (in an unrestricted market) would generate an average of $1,000,000,000 per year over their useful lives. Note, however, that because these inventions will be covered by patents, their output will be restricted during the patent term. For the reasons set forth below, the annual value represented by these inventions during the patent term will be assumed to be 75% of the amount that it would be in an (unrestricted) competitive market. The monopoly restriction and concomitant social loss imposed by the patentee thus is assumed to be equal to 25% of the social value in a competitive market. Social value is the

combination of producer and consumer surplus.

Figure 3 above represents the demand for a pioneer invention. The annual value that would be contributed by the production of the product in a competitive market is symbolized by the area DFG. We assume that the product is patented, however, so that the patentee restricts production to X, producing producer surplus represented by the rectangle EABF, and a concomitant consumer surplus DAE. The aggregate social value of the product is thus represented by the area DABF, the monopoly loss being the area AGB. Assuming a linear demand[33] for the product and constant costs as in figure 4, the monopoly loss is 25% of the total area DFG. (The triangle AGB is equal to the triangle DAE and is one half the area of the rectangle EABF.)

Year	PS (annual)	CS+PS (annual)	Monopoly loss (annual)	Value of invention	Marginal Benefit	Marginal Cost
1	$250	$750	$250	$9,000	$9,000	$250
2	$250	$750	$250	$9,000	$9,000	$500
3	$250	$750	$250	$9,000	$9,000	$750
4	$250	$750	$250	$9,000	$9,000	$1,000
5	$250	$750	$250	$9,000	$9,000	$1,250
6	$250	$750	$250	$9,000	$9,000	$1,500
7	$250	$750	$250	$9,000	$9,000	$1,750
8	$250	$750	$250	$9,000	$9,000	$2,000
9	$250	$750	$250	$9,000	$9,000	$2,250
10	$250	$750	$250	$9,000	$9,000	$2,500
11	$250	$750	$250	$9,000	$9,000	$2,750
12	$250	$750	$250	$9,000	$9,000	$3,000
13	$250	$750	$250	$9,000	$9,000	$3,250
14	$250	$750	$250	$9,000	$9,000	$3,500
15	$250	$750	$250	$9,000	$9,000	$3,750
16	$250	$750	$250	$9,000	$9,000	$4,000
17	$250	$750	$250	$9,000	$9,000	$4,250
18	$250	$750	$250	$9,000	$9,000	$4,500
19	$250	$750	$250	$9,000	$9,000	$4,750
20	$250	$750	$250	$9,000	$9,000	$5,000

Table 1

[figures in millions of dollars]

[33] The paper employs a linear model as an entry into its analysis of the patentee's situation. The use of linear analyses is common in other critiques of the patent system. See, e.g. Michael Abramowicz, *Perfecting Patent Prizes*, 56 VAND. L. REV. 127, 162-68 (2003); Douglas Gary Lichtman, *Pricing Prozac: Why the Government Should Subsidize the Purchase of Patented Pharmaceuticals*, 11 HARV. J. L. & TECH. 123, 130 (1997). Later, this assumption is dropped. See text at note 38 *infra*.

In table 1 above, we consider marginal social costs and benefits using the approach just outlined and add a number of additional assumptions. First, we continue the assumption that each year of the patent term generates new inventions whose production under competitive conditions would add $1,000,000,000 of social value (i.e., producer surplus plus consumer surplus) to the economy. And because these new inventions are patented, we continue to assume that the patentees impose monopoly restrictions on production, thus reducing the $1,000,000,000 of potential social benefit by 25% to $750,000,000. Next, we determine how to capitalize the social value of inventions. Here we make two observations. First, during the 20-year patent term, the social value is the discounted sum of the combination of producer and consumer surplus. But the social value of the invention continues beyond the end of the patent term. After the patent term the social value is the discounted sum of each year's consumer surplus (that is now enlarged when the patent rights expire). Together, these two sums equal the capitalized value of the consumer surplus unrestricted by the exercise of patent rights. Using a 5% discount rate, the value of the inventions generated by each year of patent protection would thus be 20 times earnings, or $20,000,000,000.

Alternatively, we could take a more conservative approach by drawing from the practice of investors in the securities markets. Historically, a conservative measure for the value of a stock was twelve times the annual per share earnings of the company.[34] Following our conservative approach, we capitalize only the combination of producer and consumer surplus (omitting the deadweight loss). Thus we multiply the $750,000,000 (producer plus consumer surplus) by twelve, to arrive at a capitalized social benefit of $9,000,000,000. We use this figure for social value in the analysis in table 1 above.

On these assumptions, the patent restrictions applied to the inventions which would have been produced in the absence of the patent system produce unnecessary restrictions on those inventions for 20 years. For inventions which would have been stimulated with only a one-year patent term, the actual patent term produces unnecessary restrictions for 19 years. For inventions which would have been stimulated with only a

---

[34]Investors capitalize producer surplus. Since we are concerned with social value, we capitalize the combination of producer and consumer surplus but exclude the deadweight loss.

two-year patent term, the actual patent term produces unnecessary restrictions for 18 years. Thus each year of the patent term adds restrictions on the output of a new class of inventions--i.e., those for whose generation a patent term ending a year earlier would have been sufficient--to the restrictions imposed by earlier years. The marginal social costs of patent restrictions thus rise with each year of the patent term.

We start with the most simple assumptions. Later, we will add some more complexity. Table 1 is based upon the assumption that each year of the patent term generates inventions whose aggregate capitalized value is nine billion dollars. On that assumption, the marginal benefit from each year of the patent term is a constant $9 billion. The marginal costs of the patent system, on these assumptions, gradually rise from $250 million to $5 billion. On such assumptions the patent term would have to be lengthened to 36 years before its marginal cost would rise to the level of its marginal benefits. If the patent system actually operates like the one hypothesized, there would be little room for concern that the patent term was unduly long.

I suspect that for a range each year of the patent term generates an increasing marginal return.[35] If so, over that range marginal benefit would be substantially outpacing marginal cost under the above assumptions, because new inventions would be adding their capitalized values to the computation of the marginal benefit, substantially outpacing growth in the aggregate monopoly restrictions. Ultimately, however, decreasing returns would be likely set in. And, of course, a rapid shrinkage of marginal benefit after a rapid rise would maximize the chances that marginal cost would then meet marginal benefit.

D. Ex Post and Ex Ante Values

Note that the monopoly restriction impacts society now, at the time that the output restriction occurs. The monopoly restriction is properly measured, therefore, in the present, at its full nominal dollar amount. Yet that monopoly restriction engenders an incentive to innovation only at its ex ante value. As the patent term increases in length, the difference

[35] Nordhaus appears to have assumed continually diminishing marginal returns to research investment. See Nordhaus, *supra* note 5 at 23, 73-75. Professor F. M. Scherer, commenting on Nordhaus, believes that it is more plausible to assume increasing returns to research, followed by decreasing returns. F.M. Scherer, *Nordhaus' Theory of Optimal Patent Life: A Geometric Reinterpretation*, 62 AM. ECON. REV. 422 (1972).

between the dead weight loss (measured ex post) and the incentive effect (measured ex ante) grows.

Thus, as we observed in part I, the ex ante value to the patentee of each additional year of the patent term declines over the life of the patent. This is the necessary result of discounting future returns to their present values. Because the ex ante value of the return generated by each year of the patent term declines each year of that term, then on our assumptions (all other factors remaining the same) the incentive effect of the patent term increases at a declining rate over the patent term. Each year of patent protection generates a lesser incentive to inventive activity than did the year preceding it.

				Marginal Social Benefits and Costs		
Year	CS+PS (annual)	Monopoly loss	Value of invention	Marginal Benefit	Marginal Cost	Net Marginal Benefit
1	$750	$250	$9,000	$9,000	$250	$8,750
2	$713	$238	$8,550	$8,550	$488	$8,063
3	$683	$228	$8,190	$8,190	$715	$7,475
4	$645	$215	$7,740	$7,740	$930	$6,810
5	$615	$205	$7,380	$7,380	$1,135	$6,245
6	$585	$195	$7,020	$7,020	$1,330	$5,690
7	$563	$188	$6,750	$6,750	$1,518	$5,232
8	$533	$178	$6,390	$6,390	$1,695	$4,695
9	$510	$170	$6,120	$6,120	$1,865	$4,255
10	$480	$160	$5,760	$5,760	$2,025	$3,735
11	$458	$153	$5,490	$5,490	$2,178	$3,313
12	$435	$145	$5,220	$5,220	$2,323	$2,898
13	$420	$140	$5,040	$5,040	$2,463	$2,578
14	$398	$133	$4,770	$4,770	$2,595	$2,175
15	$383	$128	$4,590	$4,590	$2,723	$1,868
16	$360	$120	$4,320	$4,320	$2,843	$1,478
17	$345	$115	$4,140	$4,140	$2,958	$1,183
18	$330	$110	$3,960	$3,960	$3,068	$893
19	$315	$105	$3,780	$3,780	$3,173	$608
20	$300	$100	$3,600	$3,600	$3,273	$328

Table 2

These considerations require that we modify the analysis in Part III. There it was assumed that each year of the patent term generates a constant amount of innovation. We now modify that analysis by discounting future returns to present values. This modification now requires us to take account of the fact that throughout the patent term, each year of patent protection generates a progressively smaller increment to the incentive to

inventive activity.

In order to discount future returns to present values, we need first to determine the discount rate. In the following example, the discount rate chosen is 5%. Let's consider an invention that generates an income stream of one dollar per year for the twenty-year patent term. Discounted to present value at the 5% rate, the income for each year falls from $1.00 in the first year to $.95 in the second year to $ .91 for the third year, and ultimately to $.40 for the twentieth year. Since a dollar of expected revenue from each year of the patent term has a present value of less than the present value of a dollar from the year preceding, we assume that the second year of patent protection generates a lesser incentive to invest in research and development than the first year. The third year generates a still lesser incentive, and so on. On this reasoning, the second, third and subsequent years generate incentives of 95%, 91%, 86%, 82%, 78%, 75%, 71%, 68%, 64%, 61%, 58%, 56%, 53%, 51%, 48%, 46%, 44%, 42% and 40% of the incentive generated by the first year revenues. Accordingly, in the following example, we assume that each of the years of the patent term provokes investment that declines in these proportions. The figures in the invention "value" column therefore are also adjusted downwards to reflect the lesser inventive activity. If the assumptions underlying table 2 accurately reflected reality, then the patent term would be almost optimal. One additional year would bring the patent term to its optimal. Beyond that, additional years would produce negative net social values.

E. Making the Model More Complex

1. Modifying the "Constant Cost" Assumption

Let's first consider dropping the constant-cost assumption employed in the preceding analysis. When marginal cost is rising, it intersects the demand curve sooner than when marginal cost is constant. Thus, in cases involving pioneer inventions in which marginal cost is rising, the dead weight loss will be less than the 25% of potential benefits of the invention assumed above in sections C and D of this Part IV. Of course, a declining marginal cost curve produces the opposite effect. A monopoly restriction in a situation of declining marginal cost produces a greater dead weight loss than in the situation of constant marginal costs. Yet economists commonly assume that a firm's short-run marginal cost curve eventually rises, because in the short run the firm is unable to adjust

the proportions at which it deploys capital with labor. As a result, its short-run marginal-cost curve takes on a U-shape.[36] Moreover, a firm with a constantly declining marginal cost curve appears to be a natural monopoly.[37] Since our focus is upon output restrictions that are generated by patents (rather than other causes), natural monopolies fall outside the domain of this paper. On the basis of these considerations, it seems appropriate now to let marginal costs either rise or remain constant. For ease of statement, let's assume that the aggregate yearly output produced under pioneer patents is produced by firms whose marginal cost curve is either flat or U-shaped.

If, for each year, most of the output value produced under pioneer patents is produced by firms with flat or U-shaped marginal cost curves, then analysis presented in sections C and D above requires adjustment. While production under conditions of flat (or constant) marginal cost and linear demand would generate the results set forth above, production under U-shaped marginal cost curves would generate lesser deadweight loss. Thus allowing production under the latter conditions into the model would increase the ratio between social benefit and social cost.

2. Including Improvement and Component Inventions

The ratio of social benefit to deadweight loss is higher in improvement inventions than in pioneer inventions. Indeed, the ratio appears to grow as the cost savings generated by the invention falls. The highest ratio of social benefit to deadweight loss is associated with modest cost improvements and the lowest ratios with inventions that approach the cost saving magnitude of a pioneer invention. Since the number of commercially valuable improvement inventions probably vastly exceeds the number of pioneer inventions, the linear model would be made somewhat more realistic by adjusting it to include improvement and component inventions. The result of this modification would necessarily increase the overall ratio of social benefit to deadweight loss generated by the patent system.

It follows that under conditions of linear demand, the ratio between marginal social benefit and marginal social cost appears to be greater than

---

[36] See, e.g., ROBERT S. PINDYCK & DANIEL L. RUBINFELD, MICROECONOMICS 224, 226 (5th ed. 2001)

[37] *Id*., at 350.

the three to one ratio considered in section D. The larger ratio is due both to the addition of improvement patents to the universe of pioneer patents first considered and to the addition of production processes involving ultimately rising (or U-shaped) marginal cost curves.

3. Dropping the Assumption of Linear Demand

Let's now drop the assumption that the demand curve is linear. Some of the current legal literature employs linear assumptions,[38] but those assumptions oversimplify reality. Probably most actual demand curves exhibit significant concavity. This would likely be true especially where the product in question appealed to a broad public and, as price dropped, became available to lower-income segments of the public. Dropping the preceding assumption of linearity, of course, substantially complicates our evaluation of the patent system's social benefits and costs. Nonlinear curves come in an endless variety of shapes and positions. In addition, there is no reason to believe that all actual demand curves are continuous rather than kinked, or indeed wrinkled, broken, or otherwise discontinuous. But bringing nonlinear demand curves into the analysis is absolutely necessary, since there is no reason to believe that actual demand curves are linear. Moreover, broadening the model may heighten our appreciation of both social problems connected with the patent system and their potential solutions. Let's take a few examples to see whether we can learn anything from them.

Many simple concave demand curves are of the form a constant over x raised to a power. Curves of this form exhibit unitary elasticity throughout their length where the exponent of x is one; inelasticity where the exponent is greater than one and elasticity where the exponent is less than one. Since monopolists maximize their profits by pricing in the elastic portions of their demand curves, demand curves of the form of k/x^2 (or k/x^3, k/x^4, etc.) do not appear interesting in a pure form, because they exhibit inelasticity throughout their length. If we add a (second) constant, however, so that the curve is of the form $k_1/x^2 + k_2$, the curve becomes elastic at higher values of x and is asymptotic to k_2. An interesting aspect of this curve is that the associated marginal revenue curve rises

[38] See, e.g. Abramowicz, *supra* note 33, at 62 68, criticizing Lichtman, *supra* note 33, at 130 (1997). Lichtman, however, recognizes the critical role that his assumptions play in his analysis. Lichtman, *supra*, at 130.

throughout. As a result, a constant marginal cost is initially higher than marginal revenue. If it is higher than the asymptote, it will never intersect with the marginal revenue curve; marginal cost would exceed marginal revenue at every level of output; and as a result there would be no output. But if a constant marginal-cost curve is lower than the asymptote, it will eventually intersect with the rising marginal-revenue curve.[39] Beyond that point, marginal revenue will exceed marginal cost for all levels of output. In such a situation, there would be no monopoly restriction on output; a monopolist in this situation would produce to capacity. A curve of this type is illustrated in figure 4 below.

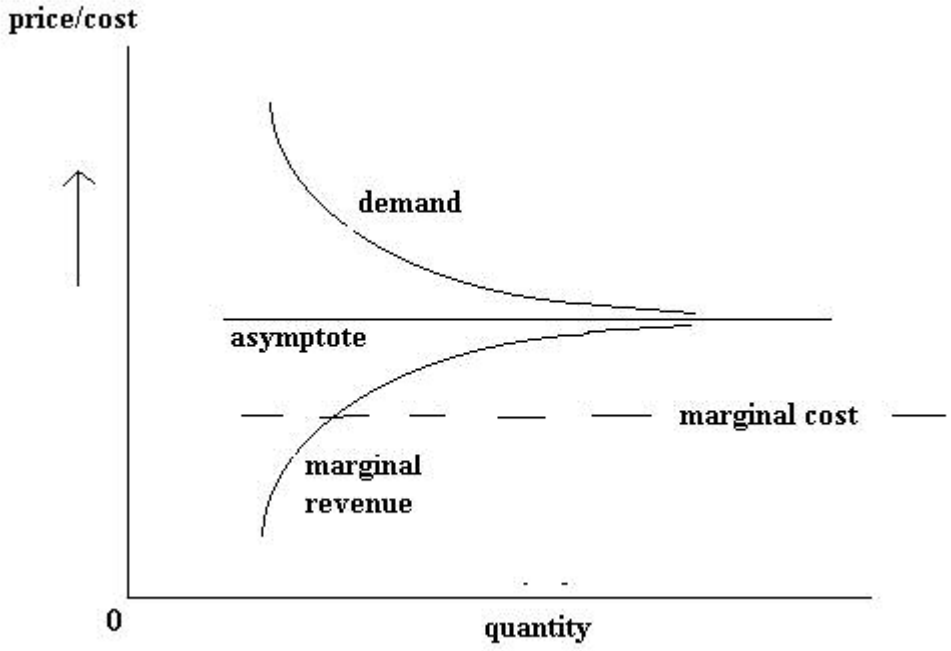

Figure 4

Now let's take a demand curve of the form of a constant over the square (or other) root of x. In figure 5 below, the demand curve is this form and the exponent is one/half, i.e., the curve takes the form of a constant over the square root of x ($k/\sqrt{x}$ or $kx^{-1/2}$). The marginal revenue curve corresponding to such a demand curve takes the form of $k/2\sqrt{x}$ or $kx^{-1/2}/2$.

[39] A U-shaped marginal cost curve would, of course, also intersect with a rising marginal cost curve.

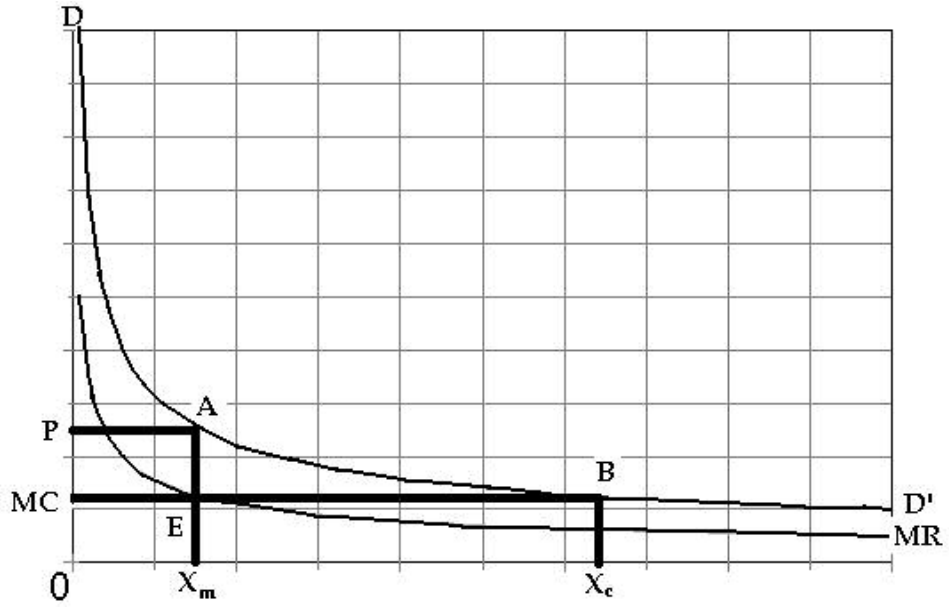

Figure 5

In the circumstances illustrated by figure 5, a pioneer inventor would favor a price/output policy determined by the intersection of its marginal cost and marginal revenue curves at point E. It would produce X_m units of output and sell them at price P. It would earn a profit represented by the area P-MC-E-A and the resulting deadweight loss would be represented by the area under the demand curve from A to B down to the marginal cost curve MC-E-B.

In this circumstance, i.e., with a demand curve of the form $k/\sqrt{x}$, it turns out that the deadweight loss is exactly equal to the seller's profits. The patent restriction thus appears more significant than it did in the case where we assumed that the demand curve took a linear form. In the former case the deadweight loss was only one half of the profits. Here it is equal to the entire amount of the profits. Still, the combination of profits and consumer surplus exceeds the deadweight loss. Let's call that combination of profits and consumer surplus total surplus for ease of presentation. Downward shifts in the cost curve would increase total surplus and (since profit equals deadweight loss) would necessarily increase the absolute amount by which total surplus exceeds deadweight loss.

The earlier projections of the balance between the social benefits and costs of the patent system made under assumptions of linear demand

do not fit these demand curves which exhibit elasticity throughout. The demand curves in this example (of the form of a constant over the square root of x) alter the ratio between benefits and costs. Inventions for which the demand takes this form would appear overall to generate higher social costs than under the earlier analysis involving linear demands. Moreover, other demand curves of the same form but involving numerically higher denominators (such as the curve generated by a constant over the square root of 2x or $k/\sqrt{2x}$) produce an even higher ratio of deadweight loss to profit. Indeed, the curve of the form $k/\sqrt{2x}$ follows the basic form of $k/\sqrt{x}$ but it is located further to the left, thereby further reducing profit in relation to deadweight loss, so that the deadweight loss produced by monopoly pricing under such a demand is actually larger than profit. Curves in the form of $k/\sqrt[3]{x}$ ($kx^{-1/3}$) or variations on them or those involving smaller negative exponents would generate even higher ratios of deadweight loss to profit. In short, the introduction of nonlinear demands shows that the potential deadweight loss generated by the patent system could be very large indeed.

4. Recapitulation and Assessment.

Let's pause to recapitulate. When we examined the social costs and benefits of a patent system under our first set of highly simplified assumptions (i.e., all patentees possessed monopoly power and all demands were linear), we found that deadweight loss was limited to an amount equal to one-half of profit and one-third of total surplus. When we broadened that model to include improvement patents, we found that the ratio of deadweight loss to social benefit would be reduced. When we examined concave demands, we found that demand curves of the form of $k_1/x2 + k_2$ would not generate any deadweight loss at all: either there would be no output (because marginal cost exceeded marginal revenue at all levels of output) or there would be output but no monopoly restriction (because marginal revenue would eventually exceed marginal cost throughout the range of possible production). Yet demand curves of different shapes produced an opposite result: When we examined concave demand curves of the form of $k/\sqrt{x}$, we found that deadweight loss was larger in relation to total surplus than under linear demands and that under many variations of that form of demand curve deadweight loss would actually exceed profit.

This examination of some of the possible shapes of demand curves thus reveals several matters: First, some types of demand curves (including both linear and some nonlinear) generate high ratios of positive welfare effects to deadweight losses. Second, there are some types of

demand curves (e.g., $k_1/x + k_2$) that are incompatible with monopoly restrictions. Third, still other types of demand curves (e.g., $k/\sqrt{x}$) would provide the context for a single-price monopolist to generate very large deadweight losses. Fourth, improvement-type patents generate a lower ratio of deadweight loss to profit under any type of demand curve; and they are likely to generate a lower ratio of deadweight loss to aggregate welfare (total surplus) so long as the production volume preceding the introduction of the improvement was sufficiently large.

Let's identify the significance of these matters for policymaking. In categories of patented inventions producing high ratios of deadweight loss to welfare (total surplus), marginal social cost meets marginal social benefit earlier than in those categories where the opposite is the case. A policy prescription seems to follow: Provide shorter patent terms for inventions with the highest ratios of deadweight loss and longer terms for inventions with lower ratios of deadweight loss. One problem that we face, however, is that while we may be able to make some judgments about how to set relative lengths of patent terms among classes of invention, we do not possess a baseline from which to set these relative terms. There is, moreover, a second problem with such a policy prescription that is discussed below. What do we know about the categories of invention that are likely to generate the highest ratios of deadweight loss to welfare?

The category of invention that is likely to produce the highest ratio of deadweight loss to welfare is likely to be a pioneer invention, as we have defined it. Thus it is a stand-alone product, rather than a component or improvement. It is likely to be a product that is desired by many people but one that many are unable or unwilling to pay the monopoly price set by the patentee. In short, the demand for the product is a highly elastic one. The demand curve may well be kinked below the monopoly price where it becomes highly elastic. Some pharmaceutical products are likely to meet this description.

In our discussion of patentee pricing producing high ratios of deadweight loss to welfare, we have been assuming that patentees set price at a single level for all purchasers. Thus our analysis has been focused upon the deadweight loss produced by a single-price monopolist. Yet is in exactly the circumstances that we have identified as giving rise to high ratios of deadweight loss to welfare that a monopolist has the greatest incentive to set a range of prices, each price geared to a different market segment. Price discrimination by a patentee is discussed below.

Here it is relevant merely to point out that patentees that are able to price discriminate among market segments may substantially reduce the deadweight losses that they would otherwise generate. The argument referred to above for reducing the patent terms of certain pioneer inventions, therefore, would not apply to these price-discriminating patentees.

V. PRICE DISCRIMINATION AND ITS BENEFITS

A. In General

Deadweight loss falls as output increases beyond the single-price monopoly output. Such increases in output can result from price discrimination. Economists recognize that price discrimination carries the potential for increasing output in monopoly markets.[40] A monopolist that practices so-called first-degree price discrimination (i.e., selling to each customer at the customer's reservation price[41]) would expand output until all customers with reservation prices above marginal cost were satisfied. In such a situation, output would be at the competitive level and there would be no deadweight loss.[42] In so-called third-degree price discrimination, a monopolist sells at its most profitable price to each of several segmented markets. It maximizes its profits when its marginal revenues from each market are the same and are equal to its marginal cost.[43] First-degree price discrimination always increases output. Third-degree price discrimination maximizes output when the demand curve in the more elastic market exhibits a greater concavity (viewed from above) than the demand curve in the less elastic market.[44]

We have observed above that a monopolist's incentive to price discriminate increases as the deadweight loss from monopoly single-pricing increases. Since price discrimination carries the potential for expanding output and reducing deadweight loss when there are

---

[40] See, e.g,, ROBERT S. PINDYCK & DANIEL L. RUBINFELD, MICROECONOMICS 372-74 (5th ed. 2001); F.M. SCHERER & DAVID ROSS, INDUSTRIAL STRUCTURE & MARKET PERFORMANCE 494-96 (3d ed. 1990).

[41] A reservation price is the highest price that a customer would be willing to pay for the product.

[42] SCHERER & ROSS, *supra* note 40 at 495.

[43] PINDYCK & RUBINFELD, *supra* note 40 at 377.

[44] SCHERER & ROSS, *supra* note 40 at 496.

substantial differences in the elasticities among markets, a patent policy that encouraged price discrimination in those circumstances would possess considerable social merit. Because patent policy is probably too crude an instrument to take into account differences in demand elasticity, a socially optimum patent policy would just endorse all price discrimination by patentees.

B. Discrimination in the Domestic Market

Despite its merits, price discrimination has not always been looked upon favorably in the United States and other nations.[45] In 1914, Congress directed section two of the Clayton Act against price discrimination that was being employed predatorily by large firms to drive their rivals from the market.[46] Later, in 1936, Congress expanded section 2 in the Robinson-Patman Act[47] in order to protect small retailers from aggressive price-cutting by chain stores who were able to secure their supplies at discriminatorily-favorable prices. In a series of legislative acts extending from 1916 to the present, Congress has sought to prevent or constrain "dumping" which is, in effect, price discrimination on an international scale.[48] Yet price discrimination is a way for a seller possessing market power both to increase its own profits as well as (under the conditions identified above) to mute the anti-social effects of its power by expanding its own output and concomitantly reducing deadweight loss. Price discrimination is widely practiced in the United States in a variety of forms.[49]

As pointed out above, the welfare loss from monopoly pricing is generally at its highest when the monopolist sells to all customers at a

---

[45] Two provisions of European Union competition law are directed against price discrimination. Article 81(1)(d) of the Treaty of Rome identifies agreements that "apply dissimilar conditions to equivalent transactions with other trading parties, thereby placing them at a competitive disadvantage" as particularly suspect. Similarly Article 82(c) identifies the application of dissimilar conditions to equivalent transactions by a dominant firm as behavior which constitutes an abuse of its position.

[46] Clayton Antitrust Act, 38 Stat. 730, ch. 323, § 2 (1914) (codified as amended at 15 U.S.C. §§ 13-13b, 21a (2000).

[47] Robinson-Patman Act, ch. 592, §§ 1-4, 49 Stat. 1526 (1936) (codified at 15 U.S.C. §§ 13-13b, 21a (2000)).

[48] Antidumping Act of 1916, ch. 463, § 801, 39 Stat. 798 (1916), Antidumping Act of 1921, ch. 14 § 201, 42 Stat. 11 (1921); Trade Act of 1974, § 321, Pub. L. no. 93-618, 88 Stat. 2043 (1974); Trade Agreements Act of 1979, 92 Stat. 162 (1979).

[49] PINDYCK & RUBINFELD, *supra* note 36 at 376.

single price. This welfare loss is aggravated where the ratio between deadweight loss and profit is high. Yet, if the monopolist were able to sell to different segments of demand at prices geared to those segments, that welfare loss might well be significantly reduced. Let's take a look at the pricing of pharmaceutical products.

To a significant degree price discrimination now appears to reduce deadweight loss in pharmaceuticals within the United States domestic market. Indeed, even at a time when price discrimination was most disfavored, Congress recognized its potential for good. Within two years after it enacted the Robinson-Patman Act to protect small business firms from large-chain-store competition, Congress enacted the Nonprofit Institutions Act in order to ensure the legality of price discrimination in favor of nonprofit institutions.[50] A major channel of distribution of pharmaceutical products involves so-called "closed door" sales to hospitals and other health-care institutions for the use of their patients.[51] In addition, pharmaceuticals are sold for a variety of prices to, or under arrangements with, wholesale drug chains, health maintenance organizations, and insurance companies.[52] A wide variety of theorists have viewed bargaining by these and similar organizations as a route for driving down pharmaceutical prices.

C. Discrimination in the Global Markets

On the international marketplace, prices often vary substantially from jurisdiction to jurisdiction, and this price variation (or discrimination) may reduce global deadweight loss. Indeed, prices of pharmaceutical products vary widely, for example, among the nations in North America and among the nations in Europe. Recent proposals to alleviate perceived high pharmaceutical prices in the United States by allowing purchases of pharmaceuticals in Canada for use within the United States have drawn attention to different pricing in different

---

[50] 15 U.S.C. § 13c (2000). The Supreme Court has nonetheless held the Robinson-Patman Act applicable to purchases by state agencies for resale in competition with retail pharmacists. Jefferson County Pharmaceutical Ass'n v. Abbot Laboratories, 460 U.S. 150, 171 (1983).

[51] Dennis S. Corgill, *Distributing Products Under the Nonprofit Institutions Act: Price Discrimination, Arbitrage, and Fraud in the Pharmaceutical Industry*, 2001 B.Y.U.L. REV. 1383, 1394-95 (2001).

[52] See, e.g., In re Brand Name Prescription Drugs Antitrust Litigation, 186 F.3d 781, 783 (7th Cir. 1999); United States v. Ferro, 252 F.3d 964, 967 (8th Cir. 2001).

national markets.[53] Pharmaceutical prices are lower in Canada than in the United States because Canada exerts control over their pricing through its Patented Medicine Prices Review Board.[54] Pharmaceutical prices in the United Kingdom are generally lower than in Germany and the Netherlands, because the U.K. government maintains an effective ceiling on their prices. Indeed, government policies on controlling pharmaceutical prices have varied substantially over the years within the European Union, giving rise to widespread arbitrage.[55]

Keying prices of pharmaceuticals to market demand in different national marketplaces would appear to be a means of both increasing the availability of these products to people that need them and to increase the profits of the pharmaceutical companies. The most obvious impediments to this approach are (1) the possibility of arbitrage diverting discounted products back to Western markets and undercutting Western prices; and (2) engendering resistance to Western pharmaceutical prices as knowledge of the discount prices provided to third-world countries spreads in the West.[56]

D. Encouraging Price Discrimination in Global Markets

It is in the economic interest of pharmaceutical manufacturers to sell their products in ways that maximize their profits. And price discrimination may further that goal. As we observed above, as the deadweight loss increases relative to the single monopoly price, price discrimination becomes ever more attractive to the seller and is likely to

---

[53] See, e.g., A. Bryan Baer, *Price Controls Through the Back Door: The Parallel Importation of Pharmaceuticals*, 9 J. INTELL. PROP. L. 109, 109-10 (2001).

[54] Patent Act, R.S.C., c. P-4 (1985)., as amended. See, e.g., Michele L. Creech, *Make a Run for the Border: Why the United States Government is Looking to the International Market for Affordable Prescription Drugs*, 15 EMORY INT'L L. REV. 593, 615 (2001) (discussing the Canadian Patented Medicine Prices Review Board); Patricia I Carter, *Federal Regulation of Pharmaceuticals in the United States and Canada*, 21 LOY. L.A. INT'L & COMP. L.J. 215, 245-49 (1999) (same).

[55] Merk v. Primecrown Ltd., (C-267/95 & 268/95) [1996] E.C.R. 6285; Merk & Co. v. Stephar BV (187/80) [1981] E.C.R. 2063; Centrafarm BV v. Sterling Drug Inc., (15/74) [1974] E.C.R. 1147.

[56] Bess-Carolina Dolmo, *Examining Global Access to Essential Pharmaceuticals in the Face of Patent Protection Rights: the South African Example*, 7 BUFF. HUM. RTS. L. REV. 137, 155 (2001) (comparing the Italian price of Fluconazole of $23.50 with the Indian price of $.95) (student note).

significantly reduce the social loss that results from a monopolistically-set single price.

Similar issues are present on the international marketplace. Large variations in wealth between the developed, developing and underdeveloped nations means that there are vast disparities in the purchasing powers of their publics. Pharmaceutical companies could benefit if they were capable of selling at a range of prices keyed to each sector of demand.[57] A major impediment to the implementation of such a program, however, is the potential for arbitrage. In Europe, where governmental interventions in markets have forced prices to comparatively low levels in certain national markets, arbitrageurs have seized the opportunities of purchasing in the low-priced markets and exporting into high-priced markets.[58] Probably a major impediment to pharmaceutical companies selling at low prices in the underdeveloped or developing world is the potential for arbitrage that such sales would engender. Arbitrageurs would be likely to purchase at third-world prices for re-export to the West for sale at North American or European prices.

Many commentators interested in increasing the availability of pharmaceuticals to third-world nations, have directed their attention to the problem of potential arbitrage, and to the potential for arbitrage to discourage low-price sales in third-world nations.[59] Most of these commentators have focused their attention on the impact of the first-sale or exhaustion doctrine and on how a doctrine of international exhaustion would facilitate arbitrage. They have also directed their attention to provisions of the World Trade Organization Agreement that prevent (or appears to prevent) governments from interfering with arbitrage operations. They have argued that that in order to effectively prevent arbitrage in such situations, governments must be enlisted in the task. Purely contractual restrictions between the exporting pharmaceutical company and its third-world customer may be inadequate, these critics have contended, to provide the needed protection. Even the customs service or health ministries of third-world nations may not be up to the

---

[57] Doctors Without Borders (Médecins Sans Frontières) reports on its web site that it has persuaded many pharmaceutical companies to sell at discount prices in third-world markets. See its report on current accomplishments at http://www.Doctorswithoutborders.org/.

[58] See note 55 *supra* and accompanying text.

[59] See, e.g., Amir Attaran, *The DOHA Declaration on the TRIPS Agreement and Public Health, Access to Pharmaceuticals, and Options under WTO Law*, 12 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 859, 879-80 (2002)

task.[60] Rather, according to these commentators, what is needed is multi-governmental cooperation: cooperation that involves the governments of the exporting nation, the importing nation and the governments of all the nations that are potential recipients of re-exports.

To what extent does the WTO regime impede arrangements that might otherwise facilitate the delivery of patented pharmaceuticals to third-world nations? This question cannot be answered without considering both the text of the Treaty, the Doha Declaration, and related developments. Considered by itself, the WTO does appear to bar the cooperation among nations that could effectively prohibit arbitraging of pharmaceutical products, as the critics have contended. Article XI of the General Agreement on Tariffs and Trade (GATT 1994) incorporated into the World Trade Organization Agreement (WTO)[61] prohibits any party to the WTO Agreement from imposing quantitative restrictions on imports or exports.[62] Accordingly, a simple reading of the literal language of Article XI would, as those commentators have suggested, appear to prohibit inter-governmental cooperation designed to prohibit the export of pharmaceuticals from poor nations or to bar their importation into wealthier ones.[63] Yet, there is much more to be said.

First, Article 31 of the TRIPS Agreement permits governments to impose compulsory licenses on patent holders in the case of a national emergency or other circumstances of extreme urgency. At the time of the anthrax scare in the United States, the U.S. government considered using this power to compel licenses on Cipro, a patented antidote to anthrax,

---

[60] *Id*., at 880.

[61] Final Act Embodying the Results of the Uruguay Round of Trade Negotiations, Done at Marrakesh, April 15, 1994, 33 I.L.M. 1125, 1154 (1994) (Annex 1A: General Agreement on Tariffs and Trade 1994)

[62] See GATT 1994, Article IX:

> No prohibitions or restrictions other than duties, taxes or other charges, whether made effective through quotas, import or export licenses or other measures, shall be instituted or maintained by any contracting party on the importation of any product of the territory of any other contracting party or on the exportation or sale for export of any product destined for the territory of any other contracting party.

[63] At least one commentator has also argued that discount sales to third-world markets would be vulnerable to attack as dumping. See Attaran, *supra* note 59, at 882. As he himself admits, however, the government of the recipient nation would be unlikely to challenge low-priced sales that benefited its own citizens, at least when there was no domestic pharmaceutical industry able to supply the domestic market and when other equally low-priced sources of the product were unavailable.

from Bayer, the German patentee.[64] So did Canada.[65] The HIV/AIDS and other epidemics in third-world nations would appear to allow them to use this Article 31 authority to impose compulsory licenses upon patented pharmaceuticals that provided needed treatments.

Article 31, however, was drafted without consideration of the fact that many poor nations lack pharmaceutical manufacturing capacity. Compulsory licensing does not help when there are no potential domestic manufacturing licensees. Accordingly, in November 2001, the ministers of the WTO member states, meeting at Doha, issued the Declaration on the TRIPS Agreement and Public Health (Doha Declaration).[66] In this Declaration, the ministers indicated that the TRIPS Agreement should be construed to allow nations to import patented pharmaceuticals from abroad to deal with national health emergencies, even though those pharmaceuticals had not been produced with the permission of the patentee. In other words, compulsory licensing would effectively extend to suppliers from abroad when the nation experiencing the health emergency lacked its own manufacturing capability. Since many of the nations experiencing HIV/AIDS health emergencies lack their own pharmaceutical manufacturing capability, the Doha Declaration was a significant corrective to the unintended rigidity of the TRIPS language. But the Doha Declaration also sheds light on the arbitraging issues discussed in this paper.

In approving the use of compulsory licenses for foreign suppliers, the ministers took steps to ensure that the entire production of the product produced under those licenses would be applied to the national health emergency and that none would be diverted to other markets. The ministers mandated that "the entire output of the relevant pharmaceuticals manufactured subject to the compulsory license should be exported to the Member in need." They indicated that the TRIPS Council should be kept informed about "the nature and quantity of pharmaceuticals being exported to a Member, the numbers of people benefiting from the solution, and the results achieved, and any evidence of diversion of products." And, recognizing that "all Members should have an obligation to ensure that the medicines in question are not diverted

---

[64] David F. Fidler, *Bioterrorism, Public Health, and International Law*, 2002 CHI. J. INT'L L. 7, 21.

[65] Susan K. Sell, *TRIPS and Access to Medicines Campaign*, 20 WIS. INT'L L.J. 481, 515 (2002).

[66] Doha Declaration on the TRIPS Agreement and Public Health. WTO Doc. WT/MIN(01)/DEC/2, 20 November 2001

from the Member's citizens for whom they were intended into other countries," they insisted upon "a commitment by all Members to take the necessary steps to prevent diversion of the relevant pharmaceuticals."

Second, it is not at all clear why contractual and licensing restrictions would not afford a degree of protection against arbitrage. The exporting pharmaceutical company could require, as a term of the sales agreement, that the purchasing firms or other organizations in the importing nation agree not to re-export and to take reasonable steps to ensure that its distributees avoid re-export. It is also possible that patents underlying these products could be employed as a base for licensing restrictions that effectively barred re-export.

Third, patent law may, in some circumstances, provide assistance in curbing or impeding arbitrage. Its usefulness depends in part upon the patent exhaustion (or first-sale) doctrine, and how that doctrine is implemented in the nations involved (or potentially involved) in arbitrage. As that doctrine is reflected in U.S. law, a patentee exhausts its rights over a particular unit of a patented product after it has sold that unit.[67] Thereafter the purchaser is generally free to resell that unit, as the purchaser pleases. This doctrine is reflected in the patent law of most other nations, producing similar results. But nations differ on how they treat sales abroad. Some nations follow a doctrine of international exhaustion, under which a sale anywhere in the world exhausts the rights of the patent holder over the units sold. The purchaser is then free to resell the product anywhere, including resales within the domestic market of the patentee. Other nations limit their use of exhaustion to their domestic markets. In these nations, a domestic sale would exhaust the patent holder's rights over units sold in the domestic market. But a sale abroad would not give the purchaser a right to resell in the domestic market. Until recently, U.S. courts have tended to apply exhaustion to unrestricted sales abroad by a U.S. patentee or a party in privity with a U.S. patentee.[68] Recently, however, the Federal Circuit has ruled that for exhaustion to apply, "the authorized first sale must have occurred under the United States patent," a view that appears to embrace a domestic, rather than

---

[67] United States v. Univis Lens Co., 316 U.S. 241, 249 (1942) ("the authorized sale of an article which is capable of use only in practicing the patent is a relinquishment of the patent monopoly with respect to the article sold.")

[68] Adams v. Burke, 84 U.S. (17 Wall.) 453 (1873); Curtiss Aeroplane & Motor Corp. v. United Aircraft Eng. Corp. , 266 Fed. 71 (2d Cir. 1920).

international, view of exhaustion.[69] Even under the traditional approach, international exhaustion was of limited scope: Thus, when foreign sales of a patented product sales have been conditioned upon their exclusion from the United States, courts have barred their importation.[70] In addition, the cases have generally refused to apply international exhaustion to the detriment of a rights-holder under a U.S. patent where the foreign sales were made without the latter's consent.[71] The Federal Circuit, however, has recently The European Union follows a policy under which sales within any member state of the Union exhaust a patent holder's rights.[72] After such a sale, the units sold may be resold anywhere within the Union. A sale outside of the Union, however, does not confer on the purchaser a right to resell within the Union.[73] Differentially-priced sales within different member-states of the Union are vulnerable to arbitrage but it is not clear whether low priced-sales of patented products outside the European Union create a potential for export back into the Union. To the extent that arbitrageurs sought to re-export pharmaceuticals to nations that followed a doctrine of international exhaustion, patent law would not provide a means for making such re-export unlawful. Although commentators have focused considerable attention on the first-sale doctrine and issues of international exhaustion, these legal issues are not necessarily the key to preventing arbitrage.

Fourth, the patent law itself contemplates the imposition of territorial limitations. While the law speaks to territorially limited assignments of rights within the United States,[74] it is clear that a patentee may grant territorially limited licenses.[75] Of course, once a licensee makes a lawful and unconditional sale to a third party, the third party can deal freely with the unit that it has purchased. If the sale takes place abroad, then the ability of the third party to export to the United States raises

---

[69] Jazz Photo Corp. v. ITC, 264 F.3d 1094, 1105 (Fed. Cir. 2001).

[70] Dickerson v. Matheson, 57 Fed. 524 (2d Cir 1893); Dickerson v. Tinling, 84 Fed. 192 (8th Cir. 1897).

[71] Boesch v. Graff, 133 U.S. 697 (1890); Griffin v. Keystone Mushroom Farm, Inc., 453 F. Supp. 1283 (E.D. Pa. 1978).

[72] Merck & Co. v. Primecrown Ltd, (C-267/95 & C-268/95) [1996] E.C.R. I-6285.

[73] See Polydor Ltd v. Harlequin Record Shops, [1982] 1 CMLR 677 (copyright case, suggestive of how EU authorities would treat interntional exhaustion issue involving patent).

[74] 35 U.S.C. § 261 (2000).

[75] See Prima Tek II L.L.C. v. A-Roo Co, 222 F.3d 1372, 1377 (Fed. Cir. 2000) ("Section 261 recognizes, and courts have long held, that an exclusive, territorial license is equivalent to an assignment . . . .")

issues of exhaustion. But if the patentee conditions the right of the licensee to sell for use solely within the jurisdiction in which the licensee is located, then a sale by the licensee not so conditioned does not convey unrestricted title to the purchaser.

In a case in which a United States patentee delivers goods to a distributor located in a particular third-world nation for distribution to users within that nation, the legal analysis would be similar. Sales to the distributor would be conditioned on resales to local users. Sales beyond the mandate of the license would be unlawful, and would not confer first-sale rights on the purchaser.

Finally, a rather obvious means for a pharmaceutical company to sell its products at low prices in a poor nation while impeding potential arbitrage would be to limit the volume of sales to estimates of local demand. The principal problem would lie in obtaining accurate estimates. But the companies' own marketing experiences both in the target market and in other similar markets may prove helpful. In addition, governments in the target markets would probably be willing to assist in the estimates of local demand. In cases where the importer was a government or government-controlled distributor, advance estimates of demand might be unnecessary. In these cases, the governmental interest would lie in ensuring that the purchased drugs would be routed to the patients who needed them, and, in order to ensure the delivery of drugs in the future, to take steps to discourage arbitrage. Patent law would support this scenario indirectly, since by ensuring that the patentee is the only source of the product, it ensures the effectiveness of the patentee's limitation of its export volume.

VI. CORRECTING THE WEAKNESSES OF THE PATENT SYSTEM

The most apparent weakness of the patent system in performing its function of fostering invention lies in the deadweight losses that this system generates. Generally, these deadweight losses are a modest price for the encouragement of invention. Indeed, since there is no deadweight loss at all without the development of both (i) a new product and (ii) one for which there is a demand, these losses are a measure of technology growth. To the extent that the exclusivity conferred on a patentee is necessary to generate an invention, the resulting deadweight loss is not a social loss at all. Yet, as discussed above, to the extent that the exclusivity

is unnecessary (as, for example, by extending longer than necessary), it can become a social cost.

The pharmaceutical industry provides an example where deadweight losses may be large when measured on a global scale. Thus it may be more of a candidate for revealing the weaknesses of the patent system than other industries. It also may be the case that the usefulness of the market in supplying information about needs to prospective inventors is at its lowest in that part of this industry that is concerned with the development of lifesaving products. It is a matter of common knowledge that cures are needed for diseases afflicting large populations. Thus the close interaction between the patent system and the pharmaceutical industry may be less socially advantageous than in other industries: in some parts of the pharmaceutical industry, market-based information is less needed; and, in that industry, the patent system's market-based incentives carry a potential for generating unduly high deadweight losses.[76]

These considerations raise the question of whether another form of financing of the development of life-saving pharmaceutical products would be desirable. Public funding of such development, if successful, would generate the larger aggregate welfare, since there would be no deadweight losses. But, to raise the question of public funding is also to raise the question of who would ultimately pay for that public funding. Well, that question, in turn, raises the question of who pays for the

---

[76] The pharmaceutical industry also reveals other dysfunctions connected with the patent system. As observed in the beginning of this article, prices of patented pharmaceuticals reflect the exclusive rights that the patent system confers upon their patentees. Consumers complain that these prices are unduly high, but exclusive patent rights are designed to produce such prices. If those prices are high, that is the result that the patent system contemplates. High prices generate the incentives necessary to stimulate inventive activity. Yet when the government responds to consumer dissatisfaction by subsidizing purchases of patented pharmaceuticals, the prices of these products will tend to rise. Government subsidization of consumers increases demand and thus price. This subsidy to consumers ultimately results in a subsidy to the pharmaceutical producers.

If, as is likely, the amended Medicare Act generates higher pharmaceutical prices in the United States, the international problem is likely to be exacerbated unless the patentees sell at discounted prices in poor countries. Purchasers in poor countries now complain that they are priced out of the market. When U.S. prices rise further, the gap between U.S. prices and affordable third-world prices will increase. As argued above, however, the gap between U.S. and third-world prices need not deprive third-world publics of patented pharmaceuticals, so long as the patentees are willing to set prices in third-world markets targeted to the demand within each market.

development of pharmaceutical products today. The answer to the latter question is that although they are paid for by the publics of Western nations, they bulk of the financial contribution comes from the American public. Prices are higher in the United States than elsewhere. Moreover, although the U.S. government does not pay list prices for pharmaceutical products, it nonetheless subsidizes them in its Medicaid program and, under recently enacted legislation, will now subsidize them in its Medicare program. In short, the U.S. public (through purchase prices, insurance premiums and/or taxes) pays a disproportionate part of the research and development cost for new pharmaceutical products now.[77]

To raise the possibility of public funding for lifesaving pharmaceutical products is also to raises the question of whether that public funding should be shared by all or most of the world's nations. The entire world benefits, or potentially benefits, from pharmaceutical research. Some formula, perhaps keyed to each nation's gross domestic product or to its per capita income might produce both a more equitably shared source of financing and a major advance in global welfare. The result, of course, would be that Western nations would make the largest contributions. That part of the result mimics the present system in which U.S. consumers pay the highest prices, Canadian and European consumers pay lower prices and the lowest are paid by consumers in those poor countries where the pharmaceutical companies sell at discount prices. This financing arrangement, however, would have the advantage of transparency, open bargaining, and an agreed-upon distribution of the burden. The extent to which other Western nations free-ride upon U.S. consumers' support of research and development would probably fall. The public funding of research would dispense with patent rights and the deadweight losses that accompany their exercise. Global welfare would advance substantially. Discontent in the third-world over patent policies that prevented or impeded their publics from treatments available elsewhere would be reduced. And a by-product of this reduced discontent would be the strengthening of the Trade Related Aspects of Intellectual Property Rights Agreement (TRIPS).

---

[77] Although U.S. consumers and taxpayers bear the burden of supporting pharmaceutical research, existing U.S. policy may confer certain advantages on American producers. See *The Pharmaceutical Industry: The Trouble with Cheap Drugs*, THE ECONOMIST, Jan. 31, 2004, at pp.59-60.