ECONOMIC THEORIES OF THE NONOBVIOUSNESS REQUIREMENT FOR PATENTABILITY: A SURVEY

by
*Vincenzo Denicolò**

In the economics literature, there have been four main approaches to the nonobviousness requirement for patentability: option value, sequential innovation, error-cost, and complementary innovation. This Article reviews these approaches and discusses their limits. All of the approaches share the premise that patenting may impose negative externalities, and thus is not always socially beneficial. When innovation is sequential, for instance, granting patent protection to trifling improvements of a path-breaking innovation may lower the pace of technological progress. Similarly, the overall incentive to innovate may be harmed if every minor component of a complex technology is separately patentable. In such circumstances, it may be desirable to deny patent protection even to genuine innovations.

I. INTRODUCTION

What is the economic function of the nonobviousness requirement in patent law?** Answering this question may help shed light on various tests of nonobviousness that have been proposed over the years, from

* This Article is a revised version of a paper presented at the conference on "Nonobviousness: The Shape of Things to Come," held at the Lewis & Clark Law School in October, 2007. I am grateful to the conference participants, and in particular my discussant Michael Katz, for helpful comments. Mailing address: Vincenzo Denicolò, Department of Economics, University of Bologna, Piazza Scaravilli, 2; I-40126 Bologna, Italy; e-mail vincenzo.denicolo@unibo.it.

** This Article is part of a multi-disciplinary conference on KSR v. Teleflex held at Lewis & Clark Law School on October 5–6, 2007. In respect for the multi-disciplinary nature of this conference, the Articles written by non-lawyers are presented herein in a modified APA citation format, rather than the usual Bluebook citation format. We have added some pinpoint citations to aid the legal reader.

"flash of creative genius" to the "teaching, suggestion, or motivation" approach recently struck down by the Supreme Court's ruling in *KSR v. Teleflex*.[1] This Article reviews several approaches to the function of nonobviousness developed in the economics literature and discusses their limits.

In assessing patentability, the most important issue faced by patent offices and the courts is, perhaps, which verifiable characteristics of the innovation should be considered to ascertain whether the nonobviousness requirement is met. Unfortunately, economists are ill equipped to address this question, since the way an innovation is conceptualized is a pre-analytical modeling choice, about which economic theory provides little guidance. By and large, the economics literature has implicitly assumed that the relevant variable for assessing nonobviousness is the "size" of the innovation. The exact meaning of size depends on the specifics of the model: it may be the magnitude of the cost reduction for process innovations, it may be the extent of the quality improvement, it may be the size of the market for the new product, but in any case it is something positively correlated to the value of the innovation. The obvious alternative to value is cost: this option, however, has proven less popular, perhaps because research and development (R&D) expenditures are notoriously hard to verify.[2]

Once the choice of how to describe an innovation is made, economists are much better equipped to analyze the optimal threshold for patentability, or at least to clarify the underlying trade-offs. In the economics literature, there have been several approaches to nonobviousness: option value, sequential innovation, error-cost, and complementary innovation. The option value approach is based on the notion of irreversible investment, and draws an analogy between the introduction of a new product or device and the exercise of an option. It argues that in a market economy there is a tendency to exercise the option too early, as today's innovators do not internalize the externality they impose on future potential innovators by depriving them of the opportunity to develop a better innovation. The nonobviousness requirement serves to protect future potential innovators against such premature inventions.

The sequential innovation approach is, in a sense, the converse of the option value approach: it views the nonobviousness requirement as a tool for protecting early innovators against competition from subsequent improvements. When innovation is sequential, such improvements cannot be achieved until and unless the basic innovation occurs, thus, protecting early innovators is of the utmost importance.

---

[1] *See* KSR Int'l Co. v. Teleflex Inc., 127 S. Ct. 1727 (2007).

[2] In addition, the U.S. Patent Act states that patentability "shall not be negatived by the manner in which the invention was made." 35 U.S.C. § 103(a) (2000). This may be interpreted as an obstacle to taking R&D costs into account in assessing patentability.

The error-cost approach views nonobviousness simply as a strengthening of the novelty requirement, designed to cope with the risk of type II errors—granting a monopoly over a technology that is already in the public domain.

Finally, the complementary innovation approach focuses on the issue of the division of profit between different firms that jointly contributed to the discovery of a complex technology. It argues that in certain circumstances denying separate patentability to some innovative components may improve the overall incentives to innovate.

While all of these approaches offer useful insights, none provides compelling policy prescriptions. All that legal scholars and practitioners can take away from existing economic analyses of the nonobviousness requirement is a clarification of certain basic, abstract questions: Why should some minimal inventive step be required on top of novelty for an innovation to be patentable? How can it be optimal to grant either full patent protection or no protection at all? These are certainly important issues, and genuine progress has been made in this area. But much work remains to be done before economic analysis can provide practical guidelines for assessing nonobviousness.

I start my survey with a discussion of the "all-or-nothing" property of current patent policy, whereby an innovation may be eligible to full patent protection or no protection at all. In Parts III to VI, I discuss each of the approaches outlined above. Part VII concludes the paper, suggesting some directions for future research.

II. AN "ALL-OR-NOTHING" SOLUTION?

Under current patent law, an obvious innovation (whatever this exactly means) cannot be patented, but as soon as the nonobviousness requirement is met the innovation enjoys a fixed level of protection, irrespective of how "nonobvious" the innovation is. How can such an "all-or-nothing" policy be optimal?

The answer many law and economics scholars have in mind is very simple. Innovations vary in size. The optimal level of patent protection, so the argument goes, is low for small innovations and increases smoothly with innovation size. Ideally, then, policy should tailor the degree of protection to each particular innovation. In practice, however, it is very difficult, if not impossible, to fine tune the policy—the well-known "one-size-fits-all" problem. The ideal curve must then be approximated with a step-wise function, implying that patent protection is granted only above a threshold innovation size, which becomes the minimal "inventive step" for patentability.

This argument is illustrated in Figure 1. The increasing curve depicts the hypothetically optimal level of patent protection as a function of the size of the innovation. As drawn, the optimal protection is smoothly increasing. The figure also depicts a step-wise approximation of the ideal

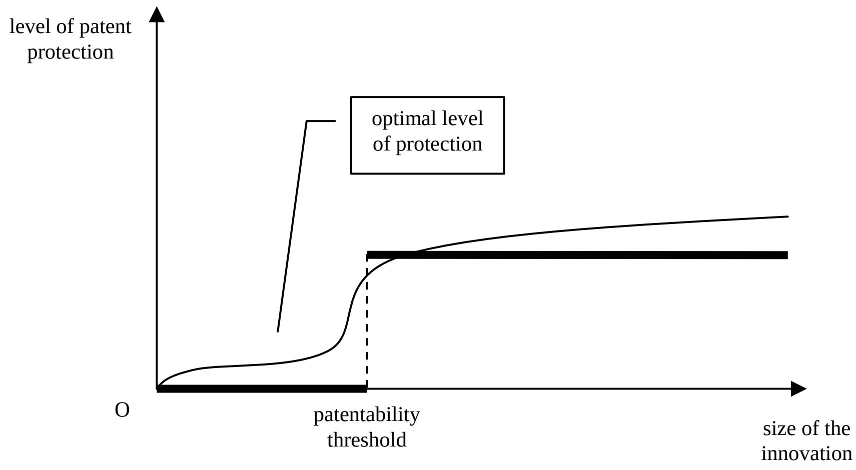

Figure 1

curve, which accounts for the constraint that patent protection can be granted or denied, but cannot be fine tuned.

Clearly, this entire argument rests on the ideal curve being increasing. But this is far from obvious. An increase in the size of an innovation generally increases the investment in research that society ought to make in that innovation. However, for any given level of patent protection, it also increases the inventor's prospective profits, automatically increasing equilibrium R&D investments. It is uncertain whether there is any need to adjust the strength of patent protection and, if so, in what direction.

The same ambiguity emerges from formal economic analyses of the optimal strength of patent protection developed since Nordhaus (1969). Nordhaus' theory suggests that the optimal level of patent protection is the outcome of a trade-off between fast innovation and static allocative distortion. Stronger patent protection entails both social benefits (innovative activity is encouraged) and costs (society bears larger monopoly deadweight losses), and the optimal level of protection must equate the marginal social benefits and costs. Both are larger for larger innovations, but it is uncertain whether the point at which equality is achieved moves to the right or to the left as the size of the innovation increases (see Figure 2 below).

Shapiro (2008) and Denicolò (2007) have shown that under reasonable simplifying assumptions, the key determinant of the optimal degree of patent protection is the elasticity of the supply of inventions—i.e., the percentage increase in the number of innovations associated with a 1 percent increase in R&D expenditures. If the elasticity is constant, the optimal level of protection is independent of innovation size. This is the case depicted in Figures 2 and 3. In Figure 2, the increasing curves represent the marginal social cost of patent protection, and the decreasing curves the marginal social benefit. The upper, solid curves represent a large innovation and the lower, dashed curves a small one. As drawn, the points of intersection of the two pairs of curves have the same abscissa.

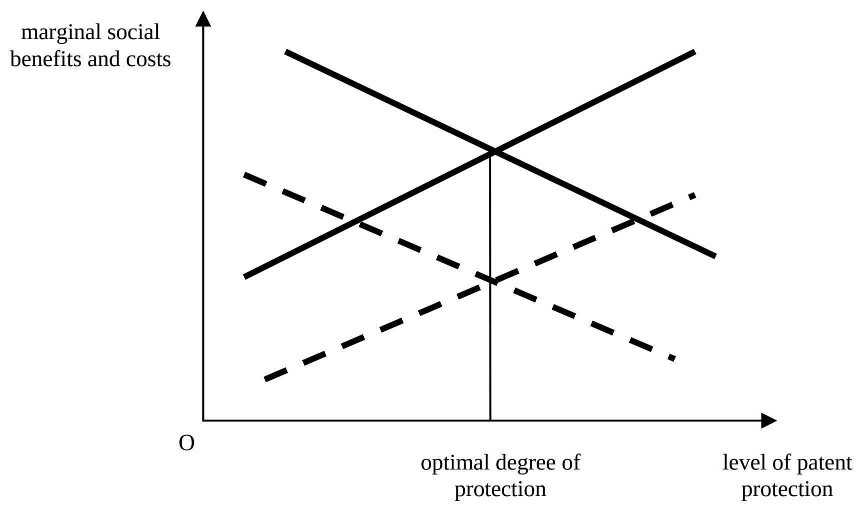

Figure 2

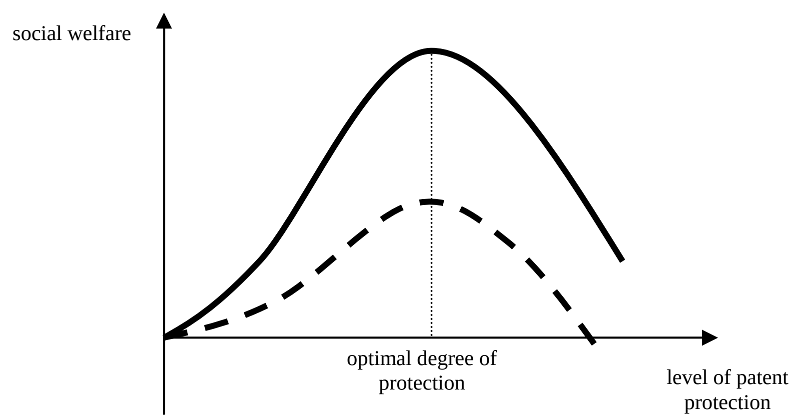

Figure 3

When the elasticity is not constant, the size of the innovation affects the optimal level of protection indirectly, as it determines the point on the "innovation production function." However, it is hard to tell whether the elasticity rises, falls, or stays constant with the level of investment in research, and hence, whether the optimal degree of patent protection increases, decreases, or stays constant with the size of the innovation.[3] The size of the innovation may also affect the optimal degree of protection through other, more indirect channels, but again, the effect is generally uncertain.[4] In sum, the claim that the ideal level of patent protection is a smoothly increasing function of innovation size does not seem to be robust.

It is always true, however, that the expected total net benefit from large innovations exceeds that from small ones. This is illustrated in Figure 3, where the dashed curve again represents a small innovation and the upper, solid curve a large one. If patents do not entail any fixed social cost, as it is implicitly assumed in Figure 3, then all innovations should be patentable and receive the same level of protection. But if patenting entails fixed social costs that are independent of the level of protection, this will shift down the social benefit curves (Figure 4). Not all innovations should now be protected. The optimal nonobviousness requirement would generally depend both on the size of the innovation (which determines the height of the gross social benefit curves depicted in Figure 3) and the magnitude of the fixed costs of patenting.

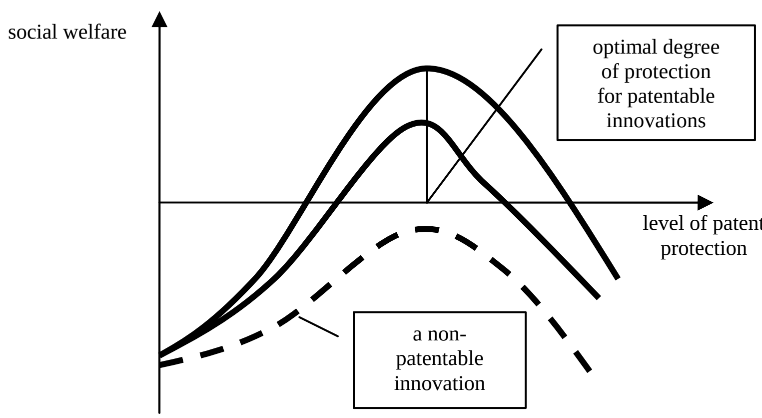

Figure 4

Thus, the all-or-nothing solution can indeed be optimal if patenting entails fixed social costs. But notice that if those costs were borne by the patent-holder, there would be no need for a nonobviousness

[3] In the standard Poisson model of patent races, the elasticity is decreasing, so the optimal degree of protection decreases with the size of the innovation. *See* Denicolò (1999). More generally, Boldrin and Levine (2005) argue that the elasticity necessarily falls for extremely large values of the investment in research. Although theirs is an asymptotical result, it may be taken to mean that the elasticity is decreasing over the relevant range, suggesting that the optimal level of patent protection actually decreases with the size of the innovation.

[4] *See*, *e.g.*, Berkowitz & Kotowitz (1982) and DeBrock (1985).

requirement.[5] If the costs exceeded the benefits, the inventor would choose not to seek patent protection, so the socially optimal decision would be implemented even if the innovation was patentable.[6] To justify a nonobviousness requirement, the fixed social cost must be a negative externality that the innovator inflicts on somebody else. Economists have identified several possible externalities of this kind, which I now turn to review.

III. THE OPTION VALUE APPROACH

The option value approach is based on the notion of irreversible investment and draws an analogy between the introduction of a new product or device and the exercise of an *option*. To illustrate this approach, Erkal and Scotchmer (2007) envision a population of potential innovators, to whom ideas to fill a market niche occur randomly. There is a constant instantaneous probability that an idea occurs to somebody, which is a key parameter of the model. Different ideas have different implementation costs, but are otherwise perfectly substitutable. Hence, if patent protection is sufficiently broad, or product market competition is sufficiently intense, after the first idea is implemented and the innovation is patented the innovation process stops; a second innovator could not recoup his innovation cost.

This simple model captures in a stylized way the notion that, in addition to opening the way to subsequent improvements, early innovations may have a pre-emptive effect. In a market economy, when deciding whether or not to invest in an idea, private inventors do not internalize the externality they are imposing on future potential innovators by depriving them of the opportunity to develop a better innovation—i.e., one with a lower implementation cost. Hence, if inventors could appropriate the value of the innovation fully, the market equilibrium would exhibit over-investment in research. That is to say, the innovation would occur too early at too high a cost.

As a result, there is room for a policy aimed at impeding premature innovations. In the Erkal and Scotchmer model, the first best solution can indeed be achieved by granting patent protection only when the implementation cost is below a critical threshold. Such a threshold is explicitly calculated by Erkal and Scotchmer, who also develop an instructive comparative statics analysis. Intuitively, if the probability that

[5] Consider, for instance, the administrative and legal costs associated with the patent application and review process. To the extent that patent offices are funded by patent fees and renewal fees, most of these costs are eventually borne by patent-holders and would-be patent-holders.

[6] This presumes that the private benefits from the innovation coincide with the social benefits. If the former fall short of the latter, the market equilibrium would be characterized by under-investment in research and under-patenting—a distortion that cannot be cured by making the innovation unpatentable.

ideas occur is large, the option value of not investing is also large, and hence the threshold should be low. Conversely, the optimal threshold is higher (i.e., it should be easier to obtain patent protection) when ideas are *scarce*.

The analysis developed by Erkal and Scotchmer provides interesting insights for policy. First, the analysis implies that nonobviousness should be assessed in terms of the R&D costs. Quite counter-intuitively, however, in the model, large R&D expenditures are a signal of obviousness. Second, patent offices and the courts should try to get a sense of the scarcity of ideas, and grant patent protection more generously when ideas are scarce.

One problem with these policy conclusions is that both R&D expenditures and the scarcity of ideas are difficult to observe, whereas policy should depend on verifiable variables.[7] In addition, the optimal policy in the Erkal and Scotchmer model could be implemented in various equivalent ways. For example, instead of using the nonobviousness requirement as a policy tool, the policymaker could use optimally chosen monetary prizes, or it could set an optimal level of patent protection, with no nonobviousness requirement at all. A richer model seems to be needed in order to compare different policy tools.

IV. THE SEQUENTIAL INNOVATION APPROACH

Although today's innovation may to some extent pre-empt future innovative activity, as stressed by the option value approach, certainly it also stimulates certain subsequent technological improvements. The notion of cumulative research, which is familiar to inventors and scholars from all fields, has indeed been the focus of a large body of economics literature on sequential innovation. This literature has highlighted the fact that each innovation builds on the previous ones and in turn is the basis for subsequent improvements. In a market economy, however, the arrival of a new innovation exerts a negative externality on past innovators, whose profits are eroded by the competition from newer and superior technologies. Such an externality can explain why certain innovations may not deserve patent protection.

There are both analogies and differences with the option value approach. In the option value approach, the negative externality is that the option of waiting for a better idea is lost; in other words, subsequent innovation is impeded. In the sequential innovation approach, by contrast, the negative externality is that a precedent innovation is discouraged. In both cases, however, there is a trade-off between present and future innovation, and the maximization of the pace of innovative

[7] Erkal and Scotchmer note that since in the model the arrival of ideas follows a Poisson process, the waiting time for the innovation is a proxy for the scarcity of ideas. However, the waiting time is also difficult to verify, since patent offices and the courts often do not know when the search for the innovation started.

activity calls for certain innovations not to be targeted. One way to achieve this goal is to make them not patentable.

There are in fact two reasons why innovators need *forward* patent protection in addition to *backward* protection when innovation is sequential: first, in the absence of forward protection future innovators could compete away the original innovators' profits; second, first innovators should be rewarded for opening the way to the subsequent improvements (Scotchmer, 1991). That is, when innovations occur in a sequence, each innovator exerts a negative externality on the previous and a positive externality on the next ones. Many authors, including Scotchmer and Green (1990), van Dijk (1995), Scotchmer (1996), O'Donoghue (1998), Denicolò (2000), and Hunt (2004), have developed the analysis of the optimal nonobviousness requirement in such a framework. Here I shall briefly review the model developed by Hunt (2004).

In Hunt's model, there is an infinitely long sequence of successive innovations. Innovative activity is sequential in that only after an innovation is achieved can the race for the next one begin. The timing of innovation is uncertain, being a probabilistic function of the amount invested in R&D. Each innovation is protected by a patent but when the next innovation arrives, the past incumbent's profits are driven to zero and the latest innovator becomes the new incumbent. The size of innovations is also uncertain. When they invest in R&D, research firms do not know how large the innovation they target will turn out to be—and hence whether it will be patentable or not—nor its per-period level of profitability.

In this world of uncertainty, the incentive to invest depends on the *expected* profits of a successful innovator. In turn, expected profits depend not only on the probability that the innovation is patentable and the expected size of the innovation, but also on the expected duration of the innovator's monopoly. A stronger patentability requirement reduces the probability that the current innovation is patentable, but it also reduces the probability that the next generation's innovation will be patentable. Thus, it prolongs the expected duration of monopoly if the innovation happens to be patentable, meaning that a stronger patentability requirement has two opposite effects on the incentive to innovate.

To see why the optimal patentability requirement is generally positive in this framework, suppose that the only objective of the policy-maker is to maximize the pace of technological progress. (If the policy-maker is also concerned with static monopoly deadweight losses, the optimal nonobviousness requirement is likely to be even stricter). Let us consider the effect of increasing the patentability threshold starting from zero. A policy move that makes very small innovations unpatentable has a negative effect on the innovator's prospective per-period profits, but this effect is second order since small innovations earn little profits even if they are patentable. The positive effect on the duration of monopoly, however, is first order, since it does not depend on the size of unpatentable innovations but only on their frequency. Thus, starting from zero a small increase in the patentability threshold unambiguously enhances the pace of technological progress.

Hunt then proceeds to a comparative statics analysis of the optimal patentability requirement. The main result is that the minimal inventive step should be higher in highly innovative industries. The intuition is that in such industries the process of profit erosion is especially rapid, so there is more need for forward patent protection.

One problem posed by this literature is that the patent system can in fact provide forward protection in two ways: the nonobviousness requirement and *leading breadth*. That is, even patentable improvements may constitute infringement on the original patent, depending on the first-generation patent's leading breadth (Green & Scotchmer, 1995; Chang, 1995; O'Donoghue, Scotchmer & Thisse, 1998). This raises various questions. Under what conditions are the nonobviousness requirement and leading breadth perfect substitutes? And, when they are not, how should forward protection be provided: through the nonobviousness requirement, leading breadth, or a combination of the two?

Denicolò and Zanchettin (2002) argue that the nonobviousness requirement and leading breadth protect early innovators in different ways. The fact that an innovation must satisfy nonobviousness requirements may block or impede subsequent improvements, thus lengthening the duration of the early innovator's monopoly. On the other hand, by finding that an improvement infringes on a basic patent, the courts can force the patent holders to bargain over profit shares, thus allowing the original innovator to capture some of the rents from the improvement. This means that the nonobviousness requirement has a "blocking effect," and leading breadth a "sharing effect."

Denicolò and Zanchettin further argue that sharing is generally better than blocking, because it does not stifle technical progress, and because the positive inter-temporal externality from the cumulative nature of innovation is at least partially internalized. As a consequence, leading breadth tends to dominate the novelty requirement as a way of providing forward protection.[8]

This means that existing economic justifications of the nonobviousness requirement as a tool to provide forward patent

[8] To be precise, leading breadth also has a blocking effect: if the second innovation is small and infringes, the profit left to the second innovator may be too low for investment to be profitable (R&D costs are sunk when bargaining between the patent holders takes place). This means that leading breadth inevitably prevents some second-generation improvements. In fact, Denicolò and Zanchettin (2002) show that while this effect is always at work, the sharing effect operates only if leading breadth is sufficiently great. Consequently, when forward protection is weak, the novelty requirement and leading breadth are perfect substitutes. By contrast, when forward protection is strong, it is preferable to use leading breadth only.

protection are at best incomplete. In addition to explaining why it may be desirable to provide forward protection, a theory must also explain in what circumstances the nonobviousness requirement is the most efficient tool to achieve that goal. This issue is still largely open. One possible reason why the nonobviousness requirement could be preferable is that leading breadth creates fragmented intellectual property rights, leading to greater transaction costs (Merges & Nelson, 1990; Heller & Eisenberg, 1998; Shapiro, 2001). However, the inclusion of those transaction costs in formal economic analyses of the nonobviousness requirement still awaits future research.

V. THE ERROR-COST APPROACH

Patents have positive breadth: if the technology space could be described as a line, for instance, a patent would typically cover an interval, not just a point. That is, patents confer a monopoly over an entire set of technologies. If the patent-holder's rights were confined literally to the device or product disclosed in the patent specification, the patent would be practically useless since it could be invented around easily. One of the major issues debated in patent design is, indeed, what determines the optimal breadth of patent protection.

The mere fact that patents have breadth, however, may create problems: even with the best of the intentions, the ambiguity of human language engenders the risk that patent offices and the courts err in defining the boundaries of the patent-holder's monopoly. There are two types of errors: denying protection over technologies actually disclosed by the innovator (a type I error), or granting a monopoly over something that is already known, and thus should stay in the public domain (a type II error).[9] The error-cost approach maintains that to reduce the risk of type II errors, patents should be granted only when the innovative technology lies sufficiently far away from the technological frontier, beyond some minimal inventive step.

Figure 5 illustrates this argument. Given that a patent must have some *lagging breadth*, but such breadth cannot be controlled perfectly by patent offices and the courts, the only way to avoid type II errors is to deny patent protection to innovations that do not constitute a sufficiently significant improvement over the prior art. This policy may be desirable even if it inevitably increases the risk of type I errors.

Here the negative externality imposed by a patent is the risk of creating deadweight losses for which there is no *quid pro quo*. Clearly, the

[9] Symmetrically, there is a risk of inventors getting a monopoly over something not yet invented, thereby stifling subsequent innovation.

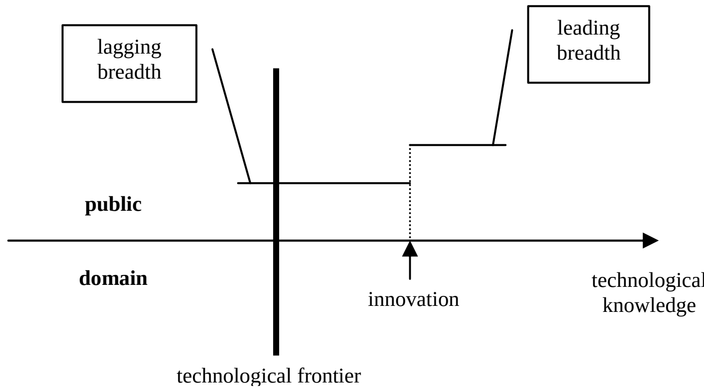

Figure 5

closer the disclosed innovative technology is to the technological frontier, the greater the risk. It follows that only sufficiently large innovations should be patentable.

Although the error-cost approach seems a natural candidate explanation for the nonobviousness requirement, to the best of my knowledge there has been no formal economic analysis along these lines so far. Many authors have noted that once a patentability threshold is somehow defined, errors may be made in the practical application of the threshold. While this is certainly true, it does not constitute a justification for the threshold itself. The error-cost approach maintains that the nonobviousness requirement is needed to minimize errors in the application of the *novelty* requirement. The implementation of the nonobviousness requirement will in turn be subject to errors, but this is another story.

Needless to say, a fully worked out theory of the nonobviousness requirement based on the error-cost approach would have to address several open issues. What exactly determines the probability of errors? Why do policymakers not simply adjust the patent's lagging breadth when the innovation is small? How large are the social costs of type II errors, and how do they compare to the costs of forestalling small innovations? It is difficult to assess the potential contribution of the error-cost approach to our understanding of the nonobviousness requirement until these and other questions are answered.

VI. THE COMPLEMENTARY INNOVATION APPROACH

The increasing complexity of modern technology and the surge in patenting have resulted in a proliferation and fragmentation of intellectual property rights. In many innovative industries, such as telecommunications, software, and biotechnology, production of new products often requires dozens, if not hundreds, of complementary innovative components, each of which may be protected by one or more patents. Since Merges and Nelson (1990), Heller and Eisenberg (1998), and Shapiro (2001), economists have recognized that the fragmentation of patent rights increases transaction costs and may lead to pricing inefficiencies. This provides another potential source of externalities that an innovator may impose on others, and hence another possible motive for denying patentability to certain innovations.

Various models of complementary innovations have been proposed in the economics literature. Some posit specialist research firms, which are capable of achieving certain innovations but not others (*e.g.*, Shapiro, 2008; Denicolò, 2007; Denicolò & Halmenschlager, 2008). Others assume that research firms are generalists, so in principle every single firm could develop all the components needed to operate the new technology (Scotchmer & Green, 1990; Ménière, 2004; Bessen & Maskin, 2008; Gilbert & Katz, 2007), although certain firms may have a comparative advantage in the search for certain components (Fershtman & Marcovich, 2006). In some models firms conduct the research sequentially, targeting one innovative component after another, while in others they invest simultaneously in all innovations.

These models differ also in the policy issues they address. Some ask very broad questions, such as whether patent protection is at all desirable. Only a few focus on the nonobviousness requirement. Specifically, Scotchmer and Green (1990) and Ménière (2004) ask whether each innovative component should be separately patentable or patent protection should be reserved for innovators who succeed in developing all the components needed to operate a new, complex technology. Denicolò and Halmenschlager (2008), by contrast, ask how strong the nonobviousness requirement should be for each separate innovative component.

Ménière (2004) builds on Scotchmer and Green (1990), differing mainly in the assumption of simultaneous rather than sequential R&D investments. Having already dealt with the sequential innovation approach in Section IV, here I focus on the Ménière paper. In his model, a strong nonobviousness requirement fosters innovation by reducing the coordination and transaction costs created by the fragmentation of intellectual property rights. However, it also reduces the incentive to disclose small innovations, thereby encouraging duplicative R&D. The optimal resolution to this trade-off requires that only those innovations that are sufficiently difficult to achieve should be separately patentable. If the probability of success is large, the innovative technology should instead be patented as a whole—that is, if a single innovator developed all the innovative components. A crucial assumption in Ménière's approach is that R&D investment is a zero-one variable, so that firms can only choose whether to invest or not. It would be interesting to extend the analysis to the case in which the probability of success is determined endogenously as a function of R&D investment.

Denicolò and Halmenschlager (2008) distinguish between two patent policy tools: the probability that an innovative component is granted patent protection (the nonobviousness requirement) and the length and breadth of protection if a patent is granted (the strength of patent protection). These two policy tools jointly determine the prospective reward to innovators, and hence the incentive to innovate. But the patentability requirements also determine the probability that the new technology reads on several patents or only one. Thus, taking the patentability requirements as a separate policy lever, they explicitly address the role of policy in determining the level of fragmentation of intellectual property rights.

Their major finding is that with complementary innovations the nonobviousness requirement should be interpreted more strictly than in the benchmark case of stand-alone innovation; that is, it should be more difficult to obtain patent protection. However, if a patent is granted, the strength of protection should be greater. The optimal expected prospective reward to innovators may be greater or lower than in the stand-alone innovation case, depending on the magnitude of the inefficiencies created by the fragmentation of intellectual property rights.

The conclusion that when the technology is complex it should be more difficult to obtain patent protection for each separate innovation component begs the question of which component should be patentable and which should not. Unfortunately, in Denicolò and Halmenschlager's model all innovative components are symmetric and hence they are effectively indistinguishable. Once again, more research is needed to draw practical prescriptions for policy.

VII. CONCLUSION

At the time of Nordhaus (1969), economists described patent policy in terms of a single variable: patent life. A lot of progress has been made since, and today there is a burgeoning literature dealing with "institutional details," like the nonobviousness requirement for patentability. In spite of this increasing interest, however, economic analysis provides little practical guidance for policy in this area. All that legal scholars and practitioners can take away from the economics literature is a clarification of the economic function served by the nonobviousness requirement. While this is crucial to a full understanding of nonobviousness, much work remains to be done before economic analysis can really succour in the assessment of patentability.

Even at the abstract theoretical level, there are several unresolved issues. One weakness of certain existing economic theories is that they artificially restrict the set of available policy tools in order to highlight the role of nonobviousness. Allowing for all conceivable tools, or even only those that are currently available in practice, makes it much more difficult to justify the existence of a nonobviousness requirement. This problem is particularly acute in the option value and the sequential innovation approaches. Another weakness is that much of the economics literature focuses on variables that are difficult to observe, whereas policy must depend on verifiable variables. Addressing these weaknesses is an important task for future research.

Having said all this, I would like to conclude with a more optimistic note. Economic theory has demonstrated that even genuine innovations can impede subsequent or precedent innovative activity, which implies that even if society's goal is to maximize the pace of technological progress, not all innovations—and *a fortiori*, not all patents—should be welcome. This nonobvious, yet unpatentable result is an important contribution to the policy debate, which can greatly help to better design our patent system.

REFERENCES

Berkowitz, M. K., & Kotowitz, Y. (1982). Patent policy in an open economy. *Canadian Journal of Economics, 15*, 1–17.

Bessen, J., & Maskin, E. (2008). Sequential innovation, patents, and imitation. *RAND Journal of Economics*, forthcoming.

Boldrin, M., & Levine, D. (2005). *IP and Market Size*. Unpublished manuscript. Retrieved March 11, 2008, from http://micheleboldrin.com/research/innovation.html.

Chang, H. (1995). Patent scope, antitrust policy, and cumulative innovation. *RAND Journal of Economics, 26*, 34–57.

DeBrock, L. M. (1985). Market structure, innovation, and optimal patent life. *Journal of Law and Economics, 28,* 223–244.

Denicolò, V. (1999). The optimal life of a patent when the timing of innovation is stochastic. *International Journal of Industrial Organization, 17*, 827–846.

Denicolò, V. (2000). Two-stage patent races and patent policy. *RAND Journal of Economics, 31*, 488–501.

Denicolò, V. (2007). Do patents over-compensate innovators? *Economic Policy, 25*, 679–729.

Denicolò, V., & Halmenschlager, C. (2008). *Complementary innovations, fragmented patents, and optimal patent policy*. Unpublished manuscript, on file with the author.

Denicolò, V., & Zanchettin, P. (2002). How should forward patent protection be provided? *International Journal of Industrial Organization, 20,* 801–827.

Erkal, N., & Scotchmer, S. (2007). *Scarcity of ideas and options to invest in R&D*. (Competition Policy Center Working Paper CPC07-080). University of California, Berkeley: CPC. Retrieved March 11, 2008, from http://repositories.cdlib.org/iber/cpc/CPC07-080.

Fershtman, C., & Markovich, S. (2006). *Patents, imitation and licensing in an asymmetric dynamic R&D race*. (CEPR Discussion Paper No. 5481). Retrieved March 11, 2008, from http://ssrn.com/abstract=902343.

Gilbert, R., & Katz, M. (2007). *Efficient division of profits from complementary innovations*. (Competition Policy Center Working Paper CPC07-076). University of California, Berkeley: CPC. Retrieved March 11, 2008, from http://repositories.cdlib.org/iber/cpc/CPC07-076.

Green, J., & Scotchmer, S. (1995). On the division of profit in sequential innovation. *RAND Journal of Economics, 26*, 20–33.

Heller, M. A., & Eisenberg, R. S. (1998, May 1). Can patents deter innovation? The anticommunist in biomedical research. *Science, 280,* 698–701.

Hunt, R. (2004). Patentability, industry structure, and innovation. *Journal of Industrial Economics, 52,* 401–425.

Ménière, Y. (2004). Non-obviousness and complementary innovations. Retrieved March 11, 2008, from http://www.cerna.ensmp.fr/Documents/YM-WP-NonObviousness.pdf.

Merges, R. P., & Nelson, R. (1990). On the complex economics of patent scope. *Columbia Law Review, 90*, 839–916.

Nordhaus, W. (1969). *Invention, growth, and welfare: A theoretical treatment of technological change*. Massachusetts: MIT Press.

O'Donoghue, T. (1998). A patentability requirement for sequential innovation. *RAND Journal of Economics, 29*, 654–679.

O'Donoghue, T., Scotchmer S., & Thisse, J. F. (1998). Patent breadth, patent life, and the pace of technological progress. *Journal of Economics and Management Strategy, 7*, 1–32.

Scotchmer, S. (1991). Standing on the shoulders of giants: Cumulative research and the patent law. *Journal of Economic Perspectives, 5*, 29–41.

Scotchmer, S. (1996). Protecting early innovators: Should second-generation products be patentable? *RAND Journal of Economics, 27*, 322–331.

Scotchmer, S., & Green, J. (1990). Novelty and disclosure in patent law. *RAND Journal of Economics, 21*, 131–146.

Shapiro, C. (2001). Navigating the patent thicket: Cross licenses, patent pools, and standard setting. In A. Jaffe, J. Lerner and S. Scott (Eds.) *Innovation Policy and the Economy* (119-150). Volume 1, Massachusetts: MIT Press.

Shapiro, C. (2008). Patent reform: Aligning reward and contribution. Retrieved March 11, 2008, from http://www.nber.org/papers/w13141. In A. Jaffe, J. Lerner and S. Scott (Eds.), *Innovation Policy and the Economy*. Volume 8, Chicago: The University of Chicago Press.

van Dijk, T. (1996). Patent height and competition in product improvements. *Journal of Industrial Economics, 44*, 151–167.