The Economics of Patents: Lessons from Recent U.S. Patent Reform

Nancy T. Gallini

New patent applications in the United States by domestic inventors climbed to nearly 150,000 per year by the late 1990s, after hovering around 60,000 per year through most of the 1980s. As shown in Figure 1, the increase in patent applications gave rise to a doubling of new patents granted per year to domestic inventors between 1985 and 1999. Every major industrial sector has been represented in this surge in activity. The high-tech sector has been most prominent, with a doubling of biotechnology patent grants and of computer software patents between 1990 and 2000 (U.S. Patent and Trademark Office, 2000). The largest 100 universities tripled their annual patent output from 1984 to 1994 (Cohen et al., 1998b), and real expenditures on research and development by small and medium-sized firms (fewer than 5,000 employees) more than doubled between 1987 and 1997 (National Science Foundation, 1997).

The last two decades have also seen important reforms in the U.S. patent system. These changes are widely perceived to have strengthened patent protection, in that patents have become easier to enforce and may be granted for a longer time. The reforms have also extended patent rights to new subject matter, such as genetically engineered life forms and business methods. The purpose of the patent reform was to stimulate innovation in the United States, which was thought to have lagged behind other industrial countries in the 1970s (Meador, 1992). Policies that strengthen and extend patent rights for the purpose of encouraging innovation find support under the conventional view of patents with its well-known public policy tradeoff: Patents provide incentives to research and to disclose information, but at the social costs of reducing the invention’s use during the patent life (Nordhaus, 1969).

Nancy T. Gallini is Professor of Economics and Research Fellow, Centre for Innovation Law and Policy, both at University of Toronto, Toronto, Ontario, Canada.
Recent research has called into question this conventional view and, consequently, the effectiveness of patents as a tool for stimulating innovation. In careful empirical studies, the value of patent protection, estimated from European patent renewal data and averaged over technology fields, has been found to be 15 to 25 percent of related research and development expenditures (Schankerman, 1998; Lanjouw, 1998). These modest estimates are consistent with findings from surveys of innovating firms suggesting that inventors rely on other ways to appropriate returns from their investment (Cohen, Nelson and Walsh, 2000). The theoretical literature shows that when research is sequential and builds upon previous discoveries, stronger patents may discourage subsequent research on valuable, but potentially infringing, follow-on inventions (Merges and Nelson, 1990; Scotchmer, 1991). In this case, the enhanced ability to enforce patents may impede rather than promote innovation, contrary to conventional belief.\footnote{See Gallini and Scotchmer (2001) for a review of the theoretical literature.}

Even if patents do not stimulate innovation, policies that promote strong patents may be justified. A second purpose of patents is to promote disclosure, a benefit that remains intact under the modern dynamic theory of patents. Inventions that would be kept secret without patents are more likely to be revealed when under patent protection, making them freely available after the patent expires. This benefit of patents must be balanced against the social costs that arise because the disclosed inventions are not freely available during the patent term: patented inventions will be used too little, may hold up subsequent research on related inventions and may generate substantial transaction costs from costly legal challenges about possible infringement. But if patents facilitate a market for technological exchange, as has been argued in the literature, licensing and other arrangements that permit the use of technology during the life of a patent may mitigate these costs.
Have the recent U.S. patent reforms been successful? That is, to what extent did the stronger patent protection cause the recent surge in patent activity? Moreover, to what extent do these patents reflect valuable inventions and the disclosure of useful information that would not have been available without the policy changes?

These questions have been central to vigorous policy debates in recent years. In a comprehensive review article, Jaffe (2000) outlines the research on these questions. Some researchers have been unable to find evidence of a causal relationship between patent reform and the rise in patents (Kortum and Lerner, 1998). Others acknowledge a link, but contend that the increase in patenting may be socially wasteful (Hall and Ziedonis, 2001), reflect inventions that are undeserving of protection (Bessen and Maskin, 2000) or impede future progress (Heller and Eisenberg, 1998). They caution that the recent U.S. patent reforms may have gone too far in protecting intellectual property rights.

With the U.S. patent reform as a frame of reference, this paper explores whether enhanced patent protection achieves the economic objectives of intellectual property protection. After providing some background on the U.S. patent reforms, I evaluate the theory and evidence on the extent to which stronger patents stimulate innovation, encourage firms to disclose their inventions and facilitate efficient technology transfer. Administrative, legal and other transaction costs that arise from extending patent protection are identified, along with proposals for alleviating them.

Recent Patent Reform in the United States

U.S. patent protection over the past two decades has been strengthened in three major ways: extending patent protection to new subject matter; giving greater power to patent holders in infringement lawsuits; and lengthening the term of patents.

With regard to the first policy change, in the 1980 decision Diamond v. Chakrabarty, the U.S. Supreme Court extended patentability to genetically engineered bacteria. In 1981, the court approved patentability of software in Diamond v. Diehr. In 1998, a federal circuit court affirmed the patentability of business methods and financial service products in State Street Bank and Trust v. Signature Financial Group. The decision validated the patent on Signature’s software system for evaluating and managing mutual funds; since then, other business methods

---

2 The Science, Technology and Economics Policy (STEP) board of the National Research Council has contributed importantly to these recent debates. See (http://ip.nationalacademies.org) for a listing of research projects sponsored by STEP.

3 The court reporter numbers for these cases are Diamond v. Chakrabarty, 447 U.S. 303, 206 U.S.P.Q. (BNA) 193 (1980); Diamond v. Diehr, 450 U.S. 175 (1981); State Street Bank and Trust v. Signature Financial Group, 149, F.3d 1368 (Fed Cir. 1998).
receiving patent protection have included Amazon.com’s one-click Internet ordering process and Priceline.com’s reverse auction method for booking airline tickets and other products on the Internet.

A second major change toward strengthening patent protection was the creation of the Court of Appeals of the Federal Circuit to handle appeals on cases involving patent infringement and validity. The creation of this specialized court was intended to bring uniformity, expertise and predictability to patent decisions. The outcome of this change has been to increase significantly the success of patent holders in legal disputes. During 1982–1990, 90 percent of patents found to be valid and infringed were upheld on appeal, as compared to 62 percent between 1953 and 1978. Decisions that a patent was invalid or that no infringement had occurred were reversed 28 percent of the time after the court’s creation, compared with the previous rate of 12 percent (Jaffe, 2000). The rate of preliminary injunctions that block the use of patented material during the infringement proceedings also has increased dramatically (Lanjouw and Lerner, 2001).

A third area of stronger patent protection involved longer patent lives for some inventions. In 1994, in compliance with the Trade-Related Aspects of Intellectual Property (TRIPs) (an agreement emanating from the Uruguay Round of the General Agreement on Tariffs and Trade), the United States adopted a 20-year patent term, starting from the date at which a patent application is filed; previously, the patent term was 17 years from the date at which the patent was issued. Ten years earlier, the Drug Price Competition and Patent Restoration Act, commonly known as the Hatch-Waxman Act, was passed. This act attempts to promote innovation in new drugs while facilitating generic entry. It restores up to five years of lost patent time spent on premarket testing and the Food and Drug Administration (FDA) approval process and formally eliminates duplicative testing of generic drugs. Moreover, the first firm to file an application for making a generic equivalent to a branded drug receives a 180-day period of exclusivity, while manufacturers of branded drugs are allowed to request a 30-month postponement of the FDA’s approval of generic drugs that arrive before their patents expire.

Potentially counterbalancing these policies that extended or strengthened patent protection are two recent policy developments. The first, the American Inventors Protection Act passed by Congress in 1999, requires that all patent applications filed in the United States and abroad be laid open for public inspection 18 months from the earliest domestic or foreign filing date. Moreover, the act

---

4 Meador (1992) documents: “It had always been one of our arguments in support of the centralized patent jurisdiction that predictability as to the validity of patents was important in promoting investment in research and production.”

5 Recently, manufacturers and generic firms have been accused of exploiting these features by entering into agreements to delay the entry of generic drugs. For example, in a complaint by the Federal Trade Commission, Abbott Laboratories reportedly offered to pay Geneva Pharmaceuticals $4.5 million per month to delay marketing its generic version of Hytrin (a hypertension drug) until the infringement litigation was resolved while retaining its period of exclusivity so that other generics could not enter. For further discussion, see (http://www.ftc.gov/opa/2000/05/hoechst.htm).
establishes a “prior use” defense against patent infringement charges for anyone who had “reduced the subject matter to practice” at least one year before the patent filing date and commercially used it prior to the filing date (Dzenitis, 2001). This ruling is generally expected to create a more certain environment for conducting research and development, since it tends to reduce the probability of a firm inadvertently working an invention that infringes part of another’s pending patent. Not everyone agrees. In a signed public letter, 26 Nobel Prize winners warned that smaller inventors may be disadvantaged in having to expose details of their inventions before the patents are issued.6

The second policy development arises from a decision in 2000 by the Court of Appeals of the Federal Circuit. In Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co., Ltd., Festo sued Shoketsu for infringing two patents on rodless cylinders. Two features of the Festo and the Shoketsu inventions were different: the magnetized material used and the number of sealing rings. The Court of Appeals of the Federal Circuit initially upheld a lower court’s decision that these differences were insubstantial and, thus, that the patent had been infringed under the “doctrine of equivalents.” This doctrine was established to prevent imitators from stealing a patented invention by making insubstantial changes.7 Eventually, the case was presented before the U.S. Supreme Court, which remanded it to the Court of Appeals for further consideration. In that round, the Court of Appeals ruled against Festo, noting that both the magnetized and coupled sealing ring features had been amended during the patent application process. In their judgment, the “doctrine of equivalents” could not be invoked for the portions of patent claims that were amended during the application process.

The Festo decision is highly controversial, and the U.S. Supreme Court has agreed to review it. Robert Bork, Festo’s attorney in the current Supreme Court case, warns that the decision would “severely undercut the protection provided by patents,” since it would allow imitators to copy the invention without retribution by changing an amended feature of the patent. At this time, it is unclear whether the Festo saga represents an attempt by the courts to rein in some of the additional protection recently afforded to patents or an oscillation in policy. The U.S. Supreme Court heard the arguments in January 2002 and will likely decide the case by spring 2002.

The effect of the more recent policies, like the American Inventors Protection Act and the Festo decision, is a question for the future. This paper focuses on the question of whether the strengthened patent protection of the 1980s and most of the 1990s inspired more innovation, disclosure and technology exchange.

---

6 See Aoki and Spiegel (1998) for a discussion and analysis of this act.
7 For a discussion of the doctrine, see Merges, Menell and Lemley (2000).
Do Stronger Patents Stimulate More Innovation?

A model of patent protection applied to a single, isolated invention predicts that stronger patents will induce more investment in research and development (Nordhaus, 1969). This model is a useful starting point, but does not accurately depict many innovation processes, particularly in high technology sectors experiencing rapid technological change. Modern models of innovation expand upon this framework by recognizing that innovation is a cumulative process that builds upon previous discoveries. More generally, they acknowledge that follow-on researchers (as well as pioneers) respond to changes in patent policy. Extending the single-invention model to incorporate these features can overturn fundamental predictions of the basic model.

To illustrate this claim, consider a policy that makes patents easier to enforce against infringement. Suppose innovation is a cumulative process that generates a continuum of improvements on some pioneer invention (Merges and Nelson, 1990; Scotchmer, 1991). When patent enforcement is strengthened, the probability increases that the researcher will prevail as the infringed against a subsequent inventor, but so does the likelihood that the researcher will lose as the infringer against a previous patent holder (O’Donoghue, Scotchmer and Thisse, 1998). If the patent is sufficiently broad such that future generations infringe previous generations of the innovation, then stronger patent rights are both a blessing and a curse. While providing the patent holder with the means to “hold up” future innovations by threatening to litigate infringers, they also increase the possibility of the patent holder being held up by previous innovators. In this setting, the link between patent strength and innovation incentives is ambiguous.

Subsequent entry can also have an impact on the effectiveness of a policy that lengthens patent life. Extending patent life may increase an entrant’s incentives to introduce an imitation during the patent period and so may not much increase the value of a patent or incentives to research (Gallini, 1992). Even if inventions can be copied only after their patents expire, incentives to innovate may decline with increases in patent life. In particular, if the inventor can develop higher quality improvements over time, then the relationship between the rate of innovation and length of a patent will have an “inverted-U” shape (Horwitz and Lai, 1996). While an increase in patent life induces the researcher to develop larger inventions, inventions occur less frequently. For patents sufficiently long, the frequency effect dominates the size effect, and so the rate of innovation declines for increases in patent life. The result that further extensions in initially long patent lives may reduce overall incentives to research may also hold if subsequent researchers, other than the pioneer, are capable of developing the improved, patentable technologies (Koo and Wright, 2001).

8 Similarly, Benoit (1985) shows that a reduction in innovation (and imitation) costs or increase in the probability of success will attract early entry by imitators, thus muting incentives to innovate.
This discussion illustrates that stronger patents may not increase the incentives to innovate. However, these incentives may be restored if the pioneer can enter into licensing contacts with subsequent inventors before they commit to any investment (Green and Scotchmer, 1995). For example, suppose a researcher is capable of improving upon a pioneer’s patented invention, but would require a license to use or market the improvement. The patent holder can increase total profits by offering a license, before the licensee engages in research and development, at a fee that respects the follower’s research costs. Such an agreement has the salutary effect of increasing both parties’ incentives to research by providing the follower with sufficient compensation to undertake the improvement, while rewarding the pioneer with a share of the rents on an invention that was inspired by the initial discovery.

The practicality of such contracts has raised a number of interrelated questions. First, identifying subsequent inventors before the pioneer commercializes the original invention may prove difficult. Second, later researchers may need to commit significant research and development investments before they can sign a useful contract with the original pioneer—but they may be reluctant to do this since those R&D costs will be sunk before the negotiations start and thus difficult to recover (Barton, 1997). Third, disparate expectations about the value of the invention may prevent the inventor and prospective licensees from settling on an acceptable royalty. Finally, transaction costs of negotiating contracts may be particularly acute for a follow-on inventor of a new product that embodies multiple patents, where each of the original patent holders can insist on a royalty or block production of the product (Merges and Nelson, 1990; Heller and Eisenberg, 1998).

To escape the costs of negotiating multiple contracts, a developer of a new product may opt for a path of less resistance. Lerner (1995) shows that firms with high costs of litigation appear to avoid research areas that are occupied by other firms, particularly when those other firms have low litigation costs.

The conventional model of patents predicts a positive and monotonic relationship between patent strength and innovation incentives. In contrast, modern theory predicts that this relationship may be nonmonotonic, depending on the relative effects of innovators being both leaders and followers and on the ease with which they can transfer their technologies. These relationships are difficult to model and test. Nonetheless, a variety of empirical studies have tackled this problem.

A useful starting point is the empirical observation that the recent extension of patents to new subject matter (biotechnology, software and business methods) in the United States increased the number of patents applied for and granted in these areas (Graham and Mowery, 2001; Lerner, 2002). A rise in the number of patents in certain areas does not ensure that the innovations are as valuable, on average, as before the change, nor that the additional protection was necessary for the development of the patented inventions. For the case of business method patents, Lerner (2002) conjectures that market incentives were adequate to spur these innovations without additional protection. If this conjecture is true, then interpreting
the rise in patent counts as additional innovation would overestimate the impact of the policy change.

Generally, patent counts are an imperfect measure of innovation (Griliches, Pakes and Hall, 1987; Lanjouw, Pakes and Putnam, 1998). For example, larger patent counts may result from a decision to seek many insignificant patents rather than a few larger patents on more fundamental inventions. Even if patent counts were an accurate measure of innovation, systematic investigations have failed to find a close connection between stronger U.S. patent policies and the overall rise in patents. For example, Kortum and Lerner (1998) note that the extension of patents to new subject matter cannot explain the overall surge in patenting; when they exclude biotech and software patents, the overall increase in patenting between 1983 and 1991 falls by only 5 percentage points, leaving over 90 percent of the increase unexplained. They also fail to find empirical support for the “friendly court” hypothesis that the creation of the Court of Appeals for the Federal Circuit led to the dramatic rise in patenting in the 1990s. If stronger patents in the United States accounted for the rise in patenting, they argued, the U.S. patent system would have become an increasingly attractive destination for foreign, as well as domestic, inventors. While patenting by U.S. inventors increased sharply both in the United States and abroad, patenting by foreigners exhibited only a modest rise in the early to mid-1990s. By a process of elimination, Kortum and Lerner attribute the enhanced patenting activity to a cause outside patent reform, specifically, to a change in the management of research and development that redirected research efforts toward more applied activities, with a modest increase in research productivity. This conclusion, however, might be tempered by the observation that the second half of the 1990s (beyond the time series of their study) reveals a pronounced increase in patent applications and patent grants in the United States by foreign inventors, lending some support for the “friendly court” hypothesis.

In an international analysis of the relationship between patent strength and innovation, Lerner (2001) examines 177 policy shifts in 60 countries over 150 years. Patent strength is measured by four features: a) whether protection existed in whole or in part for important technologies; b) the duration of the patent; c) the patent fee; and d) the existence of various limitations on patent awards (for example, compulsory licensing). The growth of patent applications by residents in the country is used as the dependent variable, while the independent variables include a dummy variable on whether the policy change is protection enhancing or reducing and the strength of protection prior to the change, among other controls. Lerner finds some support for an “inverted-U” relationship between patent strength and innovation. That is, strengthening patents has a positive effect on innovation if protection is initially low and a negative impact if patent protection is initially high. The negative part of the relationship may be attributed to features of the research process, identified in the models just discussed, that give rise to muted incentives of sequential researchers when patents are broader or easier to enforce,
or to increased imitation or the reduced frequency of inventions when patent life
is extended.9

Although it seems plausible that the strengthening of U.S. patents may have
contributed to the rise in patents over the past decade and a half, the connection
has proven difficult to verify. There are several explanations. The argument least
favorable toward patents is that they are ineffective instruments for protecting
inventions. Supporting this explanation is survey evidence that, except in a small
number of industries (pharmaceuticals, chemicals, medical equipment and some
machinery industries), patents are considered less effective relative to alternative
mechanisms for protecting intellectual assets, such as secrecy and lead time
(Cohen, Nelson and Walsh, 2000). Alternatively, the United States may be posi-
tioned at the peak of the “inverted-U” relationship, in which case further patent
protection would not spawn more innovation. The explanation most favorable
toward patents is that recent reforms may deserve some attribution for the dramatic
rise in patents (and innovation), but sufficient time has not passed to capture this
effect empirically.10

Do Stronger Patents Encourage More Disclosure?

After an invention is discovered, the researcher has the option to keep the
invention secret and risk independent discovery by another researcher, or to
disclose it through a patent grant. Patents present a bargain between society and
the inventor in which society benefits from disclosure of the invention and the
inventor receives exclusive rights over the technology for 20 years. Disclosure may
be socially valuable for several reasons. At a minimum, patents provide free access
to the inventions after the patent expires, permitting rapid diffusion after 20 years.
The information provided in patent applications may also avoid duplicated re-
search efforts by redirecting research from proven paths to those less explored.
Finally, the disclosure of ideas may inspire further new ideas. Thus, even if patents
are not the primary spur behind innovation, they may be socially worthwhile if they
induce researchers to disclose their inventions.

How valuable is the information disclosed in patent applications? According to
a survey of U.S. and Japanese firms, the information spillovers from U.S. patent
applications are relatively low (Cohen et al., 1998a). One reason is that for new,
rapidly changing technologies, patent information is largely outdated by the time
the patent is granted. Moreover, researchers may intentionally refrain from reading

---

9 International studies have also examined the relationship between patent policy and innovative
activity. Scherer and Weisburst (1995) study the introduction of pharmaceutical product patents in Italy,
and Sakakibara and Branstetter (2001) examine Japan’s conversion from a single-claim system, in which
each application could contain only a single feature for protection, to a system that allows multiple
claims.

10 See Jaffe (2000) for a thorough discussion of the impact of recent U.S. patent reform on innovation.
patent applications, since prior knowledge of inventions may expose them to the risk of treble damages in patent infringement suits.\footnote{The possibility of treble damages for willful infringement is noted in the U.S. Patent Act, 35 USCS, section 284.}

The disclosure requirement under the American Inventors Protection Act of 1999—which requires that patent applications be open for public inspection 18 months after the filing date—may improve the flow of information from patent applications. In a theoretical model of a two-stage innovation process of research and development, Aoki and Spiegel (1998) show that while the 1999 act may lead to fewer innovations and patent applications, the innovations that are discovered are more likely to be developed and to reach the product market, potentially resulting in an increase in total welfare. Their results provide an interesting twist on the conventional tradeoff of patent policy, in that the gains from implementing a smaller set of inventions potentially dominate the benefits from inventing more innovations that are less developed.

The rise in patents during the 1990s may reflect, at least in part, an increase in the “propensity to disclose” inventions in response to the patent reform, even if the policy change did not necessarily inspire these inventions. Hall and Ziedonis (2001) find support for this conjecture in the semiconductor industry, based on results of a patent production model estimated using data from 110 semiconductor firms during 1975–1996, and on comprehensive interviews with managers and executives. They attribute the increased patent propensity to the creation of the centralized Court of Appeals. By enforcing patents more strongly, they argue, the court created incentives for inventors to use litigation as a means of extracting royalties from prospective infringers. In response, firms engaged in strategic patenting, amassing large portfolios of patents for the purpose of trading them in cross-licensing agreements, when threatened by infringement suits on products they were using. The need for alternatives to litigation is particularly acute in the semiconductor industry, where products embody a large number of patented components and where a new product can provoke multiple infringement suits (Cohen, Nelson and Walsh, 2000). Hall and Ziedonis’s (2001) results also reveal that entrants specializing in the design of chips increased their patent activity for the conventional reason of protecting their inventions; however, this effect on total patenting in the semiconductor industry is dwarfed by the more prolific strategic patenting by large manufacturing firms.

Strategic patenting may be socially beneficial in encouraging the disclosure of information to other firms and in averting costly litigation. Yet the social welfare implications of this practice are controversial. The strategic accumulation of patents may redirect resources away from productive research (Hall and Ziedonis, 2001). Also, the investment of large patent portfolios required to participate in such agreements may create high barriers to entry (Barton, 1997). New semiconductor firms must spend $100 million to $200 million in licensing fees for basic technologies that may not be all that useful (Hall and Ziedonis, 2001). Of course,
this potential cost must be measured against the alternative—costly patent litigation—that is also a strong deterrent to entry.

The empirical finding that stronger protection may inspire patenting for the purpose of cross-licensing technologies points to a third incentive provided by patent policy: In addition to affecting incentives to innovate and to disclose inventions, patents may facilitate the exchange and diffusion of new technologies.

Do Stronger Patents Facilitate Technology Transfer?

After deciding to patent an invention, the patent holder can work the invention exclusively or can transfer the technology to others during the life of the patent. Technology transfer simply may be a license for the right to use the invention without the threat of a legal suit. Alternatively, it may involve the physical transfer of technologies along with valuable information, such as trade secrets or know-how, which the patent does not disclose. Components of technical know-how absent from patent applications are often transferred through licensing contracts in high technology industries, such as software, biotechnology and chemicals (Arora, 1996). Here, we examine the impact of patent strength on the decision to transfer technologies after they have been invented and disclosed.

The strength of patents may be pivotal to an innovator’s decision to license new technologies rather than to use them exclusively (Green and Scotchmer, 1995; Merges, 1998; Arora and Merges, 2000). The degree of patent protection establishes the inventor’s bargaining power in licensing negotiations, with stronger, easier-to-defend property rights generally favoring the patent holder. Empirical studies show that licensing is more prevalent in industries where effective patent protection is available, like the biotechnology and chemical industries (Arora and Fosfuri, 2000; Anand and Khanna, 2000).

More generally, strong patents promote vertical specialization (Arora, Fosfuri and Gambardella, 2001; Somaya and Teece, 2001; Tepperman, 2001). By attracting financial capital, they permit researchers to specialize in intellectual assets that are then licensed to users (Lerner, 1994). By reducing transaction costs of negotiating contractual agreements, they encourage users to license patented inputs rather than to develop their own. In the semiconductor industry, as noted earlier, firms that specialized in chip design and outsourced manufacturing tasks emerged after patent protection was strengthened (Hall and Ziedonis, 2001).

In facilitating technology exchange, patents may be self-correcting: a stronger legal right to exclude others from using an invention generally provides a stronger economic incentive to include them through licensing. This is not to suggest that patented technology will diffuse more rapidly with stronger patent protection. Under weak protection, disclosed inventions can diffuse easily through noninfringing imitations or through licenses offered to discourage costly imitation (Gallini, 1984; Gallini and Winter, 1985; Maurer and Scotchmer, 1998). Thus, the incentive
to license under strong patent protection limits, but does not eliminate, the social cost of suboptimal use of the invention during patent life.

A vivid example of how increased patent protection inspires technology transfer is given by the Bayh-Dole Act of 1980, which allowed universities to retain patent rights and to offer exclusive licenses on inventions developed with federal funds. Proponents of the legislation argued that, without exclusive licenses, private firms would be unwilling to commit investments necessary to develop university inventions beyond their “embryonic” stage. That is, they asserted that patents facilitate the use and development of pioneer inventions by other parties. The evidence is compelling that the Bayh-Dole Act has been instrumental in stimulating technology transfer between the university and private sectors (Jensen and Thursby, 2001). However, there is considerable unease within the academic community that exclusive contracts may constrain the flow of scientific knowledge that universities generate (Mazzoleni and Nelson, 1998). Other concerns include possible delays in the disclosure of scientific information and a redirection from fundamental basic research to applied research that is potentially more lucrative.

A variety of contractual arrangements, in addition to unilateral licensing, are available for transferring technologies. As already noted, cross-licensing agreements have been widely implemented in the semiconductor industry, particularly after patents became easier to enforce. In electronics, “patent pools” have become common, in which a group of patent holders agree to make their patents available to each other and to third parties at an agreed-upon royalty payment. In biotechnology, vertical integration and strategic alliances combine complementary assets of established pharmaceutical companies and start-up firms (Arora and Gambardella, 1990; Lerner and Merges, 1998). Licensing restrictions, such as grant-backs (the licensing of an innovation conditional on the licensee granting back to the licensor rights on improvements developed by the licensee), tying (licensing a patented technology with another patented or unpatented product) and exclusivity (licensing a technology to only one licensee or requiring the licensee to deal only with the licensor), are implemented across a wide range of technologies (Rey and Winter, 1998; Whinston, 2001). Research-intensive firms also enter joint ventures and other strategic alliances involving agreements about technologies to be developed in the future (Scotchmer, 1998).

While strong patents may facilitate the transfer of technology, they also may facilitate anticompetitive behavior (Chang, 1995; Gilbert and Sunshine, 1995; Anderson and Gallini, 1998; Denicolò, 2000). The U.S. Department of Justice and Federal Trade Commission’s Antitrust Guidelines for the Licensing of Intellectual Property takes a balanced view that recognizes the generally procompetitive nature of licensing (Gilbert and Tom, 2001). This view has been put to the test repeatedly since the guidelines were issued in 1995.

12 Mazzoleni and Nelson (1998) note that an exclusive license would not be necessary if there were mechanisms, such as patents, for protecting private investment made toward developing university inventions.
In cases that have involved cross-licensing, antitrust authorities typically have been permissive, as in agreements between IBM and Intel and between Hewlett-Packard and Xerox (Shapiro, 2000). Cross-licensing may remedy a situation of interlocking patents, where a patent held by one firm blocks a second firm from using its patent. Furthermore, cross-licensing provides freedom to design products more quickly without the consequence of an infringement suit (Shapiro, 2000); it may also sort firms into complementary research lines and encourage patenting on inventions that are likely to be more valuable to other firms in the agreement.

Notwithstanding these benefits, cross-licensing agreements can raise antitrust concerns, particularly when they involve a dominant firm with strong patents that are required by others. If the dominant firm threatens to withhold its technologies from those who do not agree to its terms, then antitrust enforcers face the difficult task of determining whether the refusal is anticompetitive or simply the legal exercise of intellectual property rights. For example, when Intel withdrew its intellectual property from Integraph in response to the infringement suit brought by Integraph against Intel, the Federal Trade Commission complained that Intel’s refusal to license was anticompetitive. More recently, the Court of Appeals for the Federal Circuit stated its view toward unilateral refusals involving intellectual property. In ruling on Xerox’s refusal to sell patented replacement parts to Independent Service Organizations, the Court of Appeals noted that “the patent holder may enforce the statutory right to exclude others from making, using, or selling the claimed invention free from liability under antitrust laws.” However, it acknowledged the possibility of a violation where there is “indication of illegal tying, fraud in the Patent and Trademark Office or sham litigation.”

Patent pooling has also been approved in a number of antitrust cases. For example, in 1997, the Department of Justice approved the pool of patents essential to the MPEG-2 video compression technology, which began with nine patent holders and 27 patents. Similar arrangements for pooling essential patents related to Digital Versatile Disk (DVD) technologies were approved in 1998 and 1999. Generally, patent pools have been a successful industry-established solution to conflicts over standards, particularly when each firm’s patents cover only a small component of a product. Patent pools are less likely to raise antitrust concerns if they integrate complementary components, essential to a standard, that are made available to users at reasonable royalties and if they allow members of the pool to

---

15 See Merges (1999b, c) and Shapiro (2000) for a discussion of these patent pools.
16 There are exceptions. For example, after the VL-bus standard was approved by the Video Electronics Standards Association (VESA) and became successful, Dell Computer Corporation sued some of the association’s members for violating its patent that was used in the standard. As a member of VESA, Dell had certified that it knew of no patent that the standard would violate. The FTC charged that Dell was unreasonably restraining competition and was prohibited from enforcing its patent used in the VL-bus design. For more information, see (http://www.ftc.gov/opa/1995/9511/dell.htm).
license their patented components individually as well as in a bundle with other pooled patents.

Of course, identifying patents as “complementary” or “essential” may be difficult, particularly in the early stages of an industry (Lanjouw and Lerner, 2001). The social cost of inadvertently admitting competing patents into a pool may be high. Where network externalities are present, as is typical in high-tech markets where pools tend to form, intellectual property rights can be powerful instruments for leveraging related markets (Farrell and Katz, 1998; Church and Ware, 1998).

Although stronger patents are, arguably, an impetus behind licensing and vertical specialization, the last decade has seen a wave of acquisitions, particularly in the biotechnology and information industries (Lerner and Merges, 1998; Wright, 2000; Graff, Rausser and Small, 2001). Indeed, strong patents initially promoted vertical specialization by small start-up firms in these industries. Over the past decade, however, several of these firms transferred ownership of their intellectual assets to large companies through vertical mergers. For example, in the agricultural inputs industry, dramatic consolidation occurred between 1994 and 1999 with the acquisition of top independent plant biotechnology firms and seed companies by large chemical firms, such as Calgene and Asgrow by Monsanto, Mycogen by Dow and Pioneer by DuPont (Graff, Rausser and Small, 2001).

Many arrangements for transferring intellectual property rights are private settlements of patent disputes that are made in the shadow of actual or potential litigation (Shapiro, 2000). In the formulation of policy toward licensing settlements, therefore, the expected outcome from litigation should be taken as the benchmark. Along this line, Shapiro (2001) recommends that patent settlements should not be permitted unless they generate as much consumer surplus as would be expected under a court resolution of the dispute.

What would such a rule imply for antitrust policy when patent protection is strengthened? Should antitrust authorities become more or less permissive of patent settlements in response to stronger patents? As one standard for comparison, suppose the firms decide to settle their dispute through merger rather than litigation. Since a single firm would have access to the patented innovation under either resolution—merger or litigation—the patent dispute would result in the same static outcome, ignoring court costs. However, from a dynamic point of view, the merger settlement would eliminate the stronger incentives from enhanced protection for the entrant to compete in future innovations. Similarly, a licensing agreement that transfers a share of the invention’s profits to the entrant may diminish the latter’s incentives to develop a viable substitute product. This suggests that as patent protection becomes stronger, antitrust enforcers should become increasingly vigilant in monitoring settlements that could have the effect of reducing potential competition in innovative activity. 17

Is Further Patent Reform Needed?

While the role of patent protection in stimulating innovation remains debatable, the evidence that strong patent protection encourages disclosure and technology transfer is persuasive. However, there has been no shortage of criticism of the trend toward stronger patent protection, particularly in three areas. One issue involves the consequences of extending or reaffirming patentability in the areas of software, business methods and biotechnology. A second area of controversy involves an apparent reduction in the threshold for granting patents by not requiring that patents be as “nonobvious” as has been required in the past. A third issue involves the dramatic rise in litigation over patents.

Problems with New Subject Matter

Three main categories of innovation that were not easily patentable two decades ago are now routinely patented: business methods, software and biotechnology. There has been strong criticism of extending patentability in each of these areas.

Patents on business methods were upheld in the 1998 *State Street* ruling of the Court of Appeals of the Federal Circuit, mentioned earlier. Not much time has passed since the decision, and evidence of its effect on innovation is limited, but the widely held expectation is that it will not increase innovation beyond that which would have occurred without protection. Perverse effects of the legal decision are already evident. In a study of business method patents, Lerner (2002) notes that relevant academic “prior art”—the existing set of related inventions developed by universities—has received limited acknowledgement in business method patent applications, concluding that the standard of patentability has been excessively low for this subject matter. The social cost of this error may be large, especially if the patent weapon granted to low-value business methods is used to hold up other inventions from entering the market. For example, in 1999, Amazon.com sued Barnes and Noble to block it from implementing a “one-click” online ordering process, which Amazon.com had patented. In 1998, Priceline.com asserted its patent rights on its “name your own price” reverse auction business method against Microsoft’s Expedia travel service (Hunt, 2001). Even if the courts eventually find such patents invalid, after-the-fact litigation is not a very efficient method for screening patent quality.

Prior to receiving patent protection, software was protected under copyright law. Historically, patents were reserved for physical technology, not abstract concepts such as mathematical algorithms. But over time, the distinction between the tangible and the abstract faded, and by 1981, software qualified for patent protection (Merges, 1999a). To receive patent protection, software innovations would be subjected to higher standards and more rigorous prior art searches, relative to that required for copyright protection (Merges, 1989).

Economic and legal scholars have argued strongly against the merits of this
change (Samuelson, 1990). As in the case of business methods, limited prior art for evaluating software applications may permit the issuance of undeserving patents (Graham and Mowery, 2001). Bessen and Maskin (2000) contend that the extension of patent protection to software has not been a spur to innovation in the computer industry. They argue that weaker protection would further innovation by permitting the development of complementary inventions. Open sourcing of software—providing code that individuals can use freely to develop new code or collaborate on new products—demonstrates the productive effect that complementary information sharing can have on innovation (Lerner and Tirole, 2000; Bessen, 2001).

Biotechnology, the third category with relatively newfound patent protection, has also fueled controversy. There is little disagreement that patents are an effective mechanism for protection of intellectual property in this area (Cohen, Nelson and Walsh, 2000). The controversial issues (aside from the prominent ethical issues) involve the types of innovations that merit protection. For example, genetic sequences that are isolated or purified by human intervention (such as cloned genes that produce proteins) have received both process and product patents, whereas gene fragments with undefined therapeutic value have been denied protection (Eisenberg, 1997, 1998, 2000). The fear is that granting patents on gene fragments could hold up productive research, since a firm attempting to market a product that embodied these fragments would have to negotiate licenses with multiple patent owners. Of particular concern are patents awarded on broad concepts that are vital to the success of future research but cannot be used or imitated without infringing the patents.

Starting from the premise that patent protection on business methods, software and biotechnology is not likely to be revoked, several scholars have directed their attention toward ameliorative policy amendments. The focus has often been on reducing impediments to future progress for the second-stage researcher without impinging significantly on the initial innovator’s incentives. For example, Maurer and Scotchmer (1998) advocate an “independent invention defense” (which currently exists in copyright law), in which subsequent researchers could develop a patented invention, as long as it was done independently. Such a rule may be appropriate in the area of the relatively “obvious” business methods patents. Samuelson et al. (1994) recommend that computer software should have its own unique legal protection regime, which would allow detailed and nontrivial reverse engineering. Under this regime, the initial developer would be given lead time, during which easy imitation would be prevented, while supporting valuable incremental innovation. This proposal has some precedence in the Semiconductor Chip Protection Act of 1984, which allows reverse engineering of chips to be incorporated into a new chip design, conditional on that design being sufficiently original to qualify for protection under the act. Cohen and Lemley (2000) also argue for limited reverse engineering in computer software to allow for decompilation of software that is necessary for interoperability—for programs to work with one another—so as to preserve compatibility between products. In the area of biotechnology, Eisenberg (1989) recommends an expansion of the exemption
against patent infringement claims for individuals who use research tools for experimental purposes. The current exemption is narrow, including only the use of patented inventions for satisfying scientific inquiry or philosophical curiosity; if leading to a patentable or commercialized product, it could be infringing (Samuelson and Scotchmer, 2001). A more extensive “experimental use defense,” as is embraced by several countries outside of the United States, could reduce potential holdup of future progress (Mueller, 2001). Provisions for broader experimental use have appeared in other intellectual property laws. In particular, the Plant Variety Protection Act of 1970 allows protected varieties under the act to be used in plant breeding or other research programs that may lead to new, commercially viable lines.

While these recommendations may limit the potential damage of overreaching patents after they are issued, they do not directly reduce the error and accompanying administrative and litigation costs of granting low-value patents in the first place. The gravity of this problem and remedies for alleviating it are discussed next.

A Reduction in the Standard for Nonobviousness

To merit patent protection, an invention must be nonobvious, that is, it must contribute nontrivially to prior knowledge (Merges, 1992). This rule appears to have been relaxed for some subject matter (Jaffe, 2000; Barton, 2001). The granting of a large number of questionable patents has increased the likelihood that a given invention will infringe one or more existing patents, thus provoking a barrage of litigation. Moreover, since the mid-1990s, the delay in processing patents has increased by more than 50 percent, and the backlog of applications has more than doubled (Hunt, 2001). Delays have been particularly costly in rapidly changing industries where a firm may find, only too late, that the components of a product it has developed are “owned” by one or many preexisting patent holders. “Submarine patents”—patents that emerge only after a lengthy application process—are particularly problematic when they are used to sue firms that unintentionally designed their infringing product before the submarine patent was issued (Jaffe, 2000; Shapiro, 2001).

These problems should be mitigated for patent applications filed abroad under the American Inventors Protection Act of 1999, with its rules for rapid disclosure of patent applications and its provision that prior use may be a defense against a patent infringement suit. Nevertheless, the problem of patents being granted more easily highlights a recurring theme: the same policies that are perceived to have strengthened patent rights in certain ways also have weakened them. The lower standard for patents is attributed largely to the sharp rise in applications on products and processes in new subject areas for which the U.S.

19 The policy change that set the patent term at the filing date (rather than the date of patent issuance) will likely ameliorate the problem of submarine patents, since inventors will want to expedite the process by which their patents are granted.
Patent and Trademark Office has limited expertise or access to prior art. These burdens on the patent office have compromised the quality of patents, most notably in the area of business methods.

Concerned by these problems, the U.S. Patent and Trademark Office issued the Business Methods Initiative in 2000, which recommends that more examiners be hired in the software and business methods categories; that nonpatented databases of prior art be expanded; and that a second examiner check patents on business methods. This is a start. A further step, argued persuasively by Merges (1999a), would be to invite competitors to contest the validity of an issued patent, as is done in Europe within nine months of the grant. In this case, the burden of searching prior art would shift from the patent office to the inventor’s competitors, who are likely to have better information.

While the U.S. patent process allows for the possibility of reexamining patents after they are issued, the process is nonadversarial and limited in scope and effectiveness. Consequently, it is invoked only 0.1 to 0.2 percent of the time, compared with 8 percent in Europe, and has a revocation rate of approximately one-third that of the European system (Graham et al., 2001). A more rigorous examination or opposition process would increase the quality of patents, especially in new areas where prior art is scarce. Moreover, by making firms more vigilant about their competitors’ inventions, it would also encourage the dissemination of information contained in patent applications. Finally, an opposition system could be a less costly substitute to a challenger’s alternative recourse, the legal system.

The Rise in Patent Litigation Costs

Over the past two decades, patent litigation has increased dramatically. The lucrative patent infringement award to Polaroid against Kodak in 1986 and the success of Texas Instruments in various infringement battles waged during 1985–1986 spurred other patent holders to assert their legal rights more aggressively in court (Hall and Ziedonis, 2001). The expenses of conducting a patent infringement case can cost from $1 million to several million dollars (Merges, 1999a). Lerner (1995) estimates that the costs of patent litigation started in 1991 will amount to about $1 billion (in 1991 dollars), which is 27 percent of expenditures on basic research by U.S. firms in that year. In a model of patent litigation that was estimated with patent renewal data from Germany, Lanjouw (1998) estimates that a doubling of legal fees could result in a 20 to 30 percent reduction in the mean value of patent protection in pharmaceuticals and other technologies if patent enforcement is relatively weak. While the overall litigation rate is only 1 percent of all patents, drugs and health fields are double the average (Lanjouw and Schankerman, 2001a), and in biotechnology, where patents are especially impor-

---

20 In simulations, the patent holder was assumed to have only a 50 percent chance of winning a validity challenge. For stronger patents, an increase in legal fees was estimated to have a small effect on the value of patent protection, as expected.
tant, approximately 6 percent of all biotechnology patents are litigated (Lerner, 1995).

Firms have attempted to avoid high litigation costs in various ways. One method is to direct research away from classes of projects that are occupied by firms with a comparative advantage in litigation (Lerner, 1995). Another approach noted earlier is “defensive patenting,” in which firms amass patent portfolios so that they can engage in cross-licensing when threatened by patent litigation (Hall and Ziedonis, 2001). Lanjouw and Schankerman (2001b) show that, except for pharmaceutical patents, the probability of a particular patent being involved in a litigation suit is significantly lower when the patent holder has a large portfolio of patents. Small firms have a harder time engaging in defensive patenting; as a result, the costs of litigation fall more heavily on small firms: Lerner (1995) reports that in a survey of biotech firms, 55 percent of small firms and 33 percent of large firms report that litigation is a deterrent to innovation. Moreover, small firms are disadvantaged by costly preliminary injunctions: Firms requesting injunctions tend to be twice as large as those that do not and are significantly larger than the defendant (Lanjouw and Lerner, 2001).

But stronger patents do not always tilt the balance against smaller entities. They can also bring small firms clear bargaining power, even against larger firms, and also help small firms attract financing. Moreover, two remedies available—the prior use and experimental use defenses (although limited)—may give some protection to small firms that would otherwise face infringement challenges. Where these remedies are deficient, Lanjouw and Schankerman (2001a) propose a market for patent litigation insurance to help reduce the relatively high burden of patent disputes to small firms. Private insurance schemes for patent litigation, currently available in the United States, are limited in coverage and available only at high premiums, owing to the lack of information in assessing risks. In their empirical analysis of litigated patents, Lanjouw and Schankerman estimate a relationship between the probability of lawsuits and observable characteristics of patents and their owners that, they suggest, could be used to identify risks of patent litigation and to allow for the design of more efficient insurance schemes.

Conclusion

It may be premature to pass judgment on the policy decisions that appear to have strengthened U.S. patent protection in the 1980s and 1990s (Jaffe, 2000). Perhaps sufficient time has not passed for firms to respond to the policy changes; perhaps the results are confounded by other changes that are difficult to isolate from the policy effect; or perhaps appropriate attribution has not been given to indirect effects of the policy changes. For example, alliances between research firms, facilitated by strong property rights, may have allowed innovators to take advantage of new opportunities or to redirect their research toward more productive uses.
As economists and legal experts continue to uncover the mysteries of the intellectual property system, current policy should evolve to reflect what we do know with confidence. We know that patents may simply be too blunt an instrument to apply to all technologies; so adopting specialized intellectual property regimes for certain subject matter may be warranted. We also know that the U.S. Patent Office has become overburdened with applications in areas for which it has limited expertise or prior art, resulting in a reduction in patent standards and a rise in litigation. We suspect that these problems would likely be alleviated under an opposition procedure, which engages private parties. Broadening the infringement exemption for experimental use of patented inventions and permitting the exchange and restructuring of intellectual property rights between research firms, consistent with antitrust laws, are also sensible proposals for improving the research climate.

As new technologies emerge, so will patent, legal and antitrust rules that govern the granting, enforcement and exercise of intellectual property protection. In designing effective rules, we can no longer rely on the simple tradeoff—that patents stimulate innovative activity but at the cost of constraining use of the innovation output—as our guide. In an environment of cumulative innovation, patents can undermine protection on the very inventions they seek to protect. Patent reform affects dimensions of the innovative process that stretch far beyond the incentives to innovate. Policies that extend and strengthen intellectual property rights can have profound effects on the direction of technological change, on strategies used by firms to manage their intellectual assets, on industrial restructuring and on the very institutions that establish and enforce intellectual property rights.

I am grateful to Ashish Arora, Jim Bessen, Wes Cohen, Brad De Long, Rich Gilbert, Bronwyn Hall, Alan Krueger, Jenny Lanjouw, Josh Lerner, Peter Menell, Rob Merges, Pam Samuelson, Suzanne Scotchmer, Alex Stack, Andrew Tepperman, Michael Waldman, Ralph Winter and Brian Wright for their many useful comments. Timothy Taylor deserves special thanks for his valuable insights and skillful editing of multiple drafts of the paper. Financial support was generously provided by the Social Science and Humanities Research Council of Canada.

References


Somaya, Deepak and David Teece. 2001.


