

Probabilistic Patents

Mark A. Lemley and Carl Shapiro

For many years, economists typically conceptualized patents as well-defined property rights giving their owners either a monopoly over some market or at least a significant competitive advantage in that market due to control over a product improvement or a low-cost method of production (Nordhaus, 1969; Reinganum, 1989). Once a patent was issued, this approach tended to assume that the patent was valid, that it granted a right of definite scope, and that users of the patented technology respected that right or were forced by courts to do so. Treating patents as well-defined rights to exclude rivals has permitted economists to focus on important and complex relationships among patents, innovation, competition and the diffusion of technology.

More recently, however, scholars and policymakers have begun to look more closely at the empirical evidence regarding the issuance of patents and patent litigation. Nearly 200,000 patents are issued every year after a very limited examination process. Most issued patents turn out to have little or no commercial significance, which is one reason that only 1.5 percent of patents are ever litigated, and only 0.1 percent of patents are ever litigated to trial. Given these uncertainties, economists have increasingly recognized that a patent does not confer upon its owner the right to exclude but rather a right to *try* to exclude by asserting the patent in court (Shapiro, 2003a). When a patent holder asserts its patent against an alleged infringer, the patent holder is rolling the dice. If the patent is found invalid, the property right will have evaporated.

■ *Mark A. Lemley is the William H. Neukom Professor of Law, Stanford Law School, Stanford, California, and of counsel, Keker & Van Nest LLP, San Francisco, California. Carl Shapiro is the Transamerica Professor of Business Strategy, Haas School of Business, and Professor of Economics, University of California at Berkeley, Berkeley, California.*

The risk that a patent will be declared invalid is substantial. Roughly half of all litigated patents are found to be invalid, including some of great commercial significance. For example, Chiron's patent on monoclonal antibodies specific to breast cancer antigens was invalidated by a jury in 2002 in a suit in which Chiron had sought over \$1 billion in damages from Genentech. A decision by the U.S. Court of Appeals for the Federal Circuit to invalidate an Eli Lilly patent on Prozac in 2000, less than two years before the patent was set to expire, caused Lilly's stock price to drop 31 percent in a day.

Virtually all property rights contain some element of uncertainty. The owner of real estate may find that the title to that property is flawed; title insurance exists to deal with this risk. The (careless) owner of a trademark may find that its mark has been used so widely as to become a generic term, thus losing trademark protection. But the uncertainty associated with patents is especially striking, and indeed is fundamental to understanding the effects of patents on innovation and competition. There are two fundamental dimensions of uncertainty: 1) uncertainty about the commercial significance of the invention being patented, and 2) uncertainty about the validity and scope of the legal right being granted. Uncertainty about commercial significance is critical when studying the process by which patents are issued. Uncertainty about validity and scope are critical when studying the enforcement and litigation of patents.

This article explores the economics of probabilistic patents. We begin with a brief description of the system by which patents are issued and litigated in the United States. We then discuss uncertainty regarding commercial significance and how patent applicants and would-be licensees hedge against that uncertainty. We turn next to proposals to reform the system by strengthening the examination process through which patents are granted. This literature emphasizes that expending the resources necessary to increase the certainty of issued patents across-the-board may not make economic sense given the vast number of patent applications and the small number that end up being commercially important. More effective are patent reforms that focus resources on patent applications that are likely to be of commercial significance.

We then explore the incentives of patent holders and alleged infringers to settle their disputes rather than litigate them to completion. Indeed, virtually every patent licensing and cross-licensing agreement can be seen as the settlement of a patent dispute. However, the frequency or form of such private settlements may not serve the public interest, because litigating patent disputes to completion tends to generate positive externalities, by clarifying the limits of patent protection if the patent is upheld or encouraging wider use of the innovation if the patent is invalidated. While settlements of patent litigation between actual or potential rivals are normal and generally desirable, they also are agreements between competitors that can limit competition. In fact, a recent flurry of court cases have recognized the impact of uncertainty while scrutinizing the antitrust implications of patent settlements.

Patent Prosecution and Patent Litigation in the United States

Patenting is big business in the United States and throughout the world.¹ Inventors file over 350,000 patent applications a year with the U.S. Patent and Trademark Office (PTO), a number that has grown steadily (U.S. PTO Annual Report, 2003), and spend over \$5 billion a year just on the process of obtaining those patents (Lemley, 2001). The PTO grants nearly 200,000 new patents a year, a number that has roughly doubled over the past 15 years (U.S. PTO Annual Report, 2003, Table 6).

Patents are rewards for those who have contributed to economic growth through their inventions. Any resulting market power enjoyed by a patent holder is typically considered a social cost that is necessary to stimulate innovation and provide a return on R&D expenditures. When patents are granted covering technologies that were already known or were obvious, the resulting patents cause social costs without offsetting benefits. Therefore, the lower the quality of patents—that is, the greater is the fraction of patents improperly issued—the less efficient the patent system is at stimulating innovation. Complaints about patent quality are not new, but they have grown louder in recent years (Federal Trade Commission (FTC), 2003; National Academies of Science (NAS), 2004).

Patent Applications and Patent Prosecution

The scope of a patent is defined by its “claims.” Examiners at the Patent and Trademark Office look at the claims in each patent application to see if the invention described in the claims meets the statutory requirements for patentability; the key requirements are that the invention be novel and not obvious. The PTO tests for novelty by looking for “prior art,” typically as described in other patents, in publications and in existing products. The applicant is required to disclose relevant prior art of which it is aware, but not to conduct a thorough search for prior art. The PTO also grants a patent only if it believes that the invention would not have been obvious “to a person of ordinary skill in the art” at the time the application was filed.

The applicant is required to disclose its invention in the patent application, which will then be made public when the patent is issued (or 18 months after the application is filed if the applicant also seeks patents outside the United States, as most do). The disclosure must be sufficient so that those skilled in the art can make and use the invention without undue experimentation.

Inventors regularly file patent applications without any clear idea of whether the invention will be a commercial success and in some instances whether the category of invention is even patentable at all. In part, this dynamic is driven by the strong incentives to file applications early. Inventors who commercialize a product,

¹ More detailed descriptions of the operation of the U.S. patent system are available from many sources. See, for example, National Academies of Science (2004, Appendix A, “A Patent Primer”) or Schechter and Thomas (2003).

publish a paper or disclose an idea to the public have only a year in the United States to get a patent application on file (35 U.S.C. § 102(b)), or the innovation passes into the public domain. Those who wish protection abroad have even stronger incentives to file quickly, both because Europe lacks a one-year grace period and because if more than one inventor claims the same invention, every country but the United States will give the patent to the first person to *file* an application, rather than the first to *invent*. In theory, these policies encourage prompt filing of patents and increase the disclosure of ideas. They also help reduce the risk that firms will make investments in good faith and only find out later that they have infringed on the patent of another firm.² But the incentives for early filing also mean that many inventors file patent applications before they have any good sense of whether their inventions have commercial significance. For example, inventors in the fields of pharmaceuticals and biotechnology file applications on a large number of drugs or therapies before they know whether those drugs will be safe and effective.

Moreover, the scope of patentable subject matter has been expanding over time (Gallini, 2002). The Federal Circuit Court of Appeals added software to the list of patentable inventions in the 1980s and 1990s, and “business methods” in 1998.³ Because the scope of patentable subject matter can change over time, inventors may file patent applications even in areas that are not currently eligible for patent protection, simply to hedge their bets.

Once an application is filed, the patent applicant negotiates with the patent examiner over the allowability and scope of the claims. The burden is on the Patent and Trademark Office to provide a reason *not* to issue a patent sought by an applicant. If the applicant is dissatisfied with the claims allowed by the patent examiner, the applicant can file a continuation application even after receiving a patent and thus continue to seek a patent with broader claims. The applicant can add new matter to the continuation application and claim the invention with additional, new elements, albeit with a later date of invention. Remarkably, there is no limit to the number of times the patentee can seek continuation of a patent application (Lemley and Moore, 2004). Applicants are even allowed to amend their

² Prior to 1999, U.S. patent applications were not disclosed prior to the issuance of the patent. Some patent applications remained hidden from public view for years, even decades, while the PTO considered those applications. These so-called “submarine patents” allowed patent holders to engage in significant opportunism, in part by strategically slowing down the PTO process and in part by amending claims to capture new products introduced into the market well after the patent application was initially filed. Submarine patents are less of a concern since the law was changed in 1999 to require the disclosure of most U.S. patent applications 18 months after they are filed.

³ Perhaps the best-known patent on a “business method” is Amazon’s 1999 patent no. 5,960,411. This “one-click” patent protects Amazon’s online ordering system, which allows consumers to make purchases with a single mouse click. Amazon generated a great deal of attention by asserting this patent against its most direct online rival, Barnes & Noble. For more background, see Lerner (2002) and his discussion of the case of *State Street Bank & Trust Co. v. Signature Financial Group, Inc.* (149 F.3d 1368 [1998]), in which the Federal Circuit explicitly rejected the notion that “business methods” could not be patented. The *State Street* case involved a software program that was used to value mutual funds.

applications to capture products that are appearing in the market, so long as they (arguably) stay within the bounds of the invention described in the initial application, which can be broad and rather vague. Furthermore, applicants who are not content with the examiner's decision have appeal rights.

While the examination process at the Patent and Trademark Office takes nearly three years on average (Allison and Lemley, 2000), a patent examiner spends only 18 hours per application on average during those three years reading the application, searching for and reading prior art, comparing the prior art to the application, writing one or more provisional rejections, reviewing responses and amendments, often conducting an interview with the applicant's attorney and writing a notice of allowance (Lemley, 2001; FTC, 2003). Even with this quick look, the PTO recently reported a backlog of more than 750,000 patent applications (U.S. PTO Annual Report, 2003, Table 5). Further, legal scholars who have studied the patent prosecution process have pointed to structural problems that encourage the PTO to grant patents of doubtful quality, including high examiner turnover and an incentive system that rewards examiners for allowing but not for rejecting applications (Merges, 1999; Thomas, 2001). As a result, the overwhelming majority of patent applications in the United States, perhaps 85 percent, ultimately result in an issued patent—far more than in Europe and Japan (Quillen, Webster and Eichman, 2003; NAS, 2004).⁴

Patent Litigation and Damages

Patent lawsuits take place in the federal courts, usually in front of a jury. Appeals of patent decisions go to the Federal Circuit Court of Appeals, a specialized appeals court for patent cases established in the early 1980s. Many observers believe that the creation of the Federal Circuit has led to decisions more favorable to patent holders (Dreyfuss, 1989; Kortum and Lerner, 1999).

The vast majority of patents are never asserted in litigation. Only 1.5 percent of all patents are ever litigated, and only 0.1 percent are litigated to trial (Lanjouw and Schankerman, 2001; Lemley, 2001), though litigation rates vary by industry and reach as high as 6 percent in biotechnology (Lerner, 1995). Great care must therefore be taken when interpreting data from any sample of litigated patent cases. The patents involved in litigation are those that are important enough commercially to justify the costs of litigation and for which the parties were unable to reach a mutually attractive settlement.

⁴ One cannot simply divide the number of issued patents into the number of applications to obtain the grant rate, for several reasons. First, because the number of applications increases from year to year, and because it takes almost three years on average for the PTO to issue a patent, the proper comparison would be between the number of patents issued in a given year and the number of applications filed three years earlier. Second, over a quarter of all U.S. patent applications are so-called "continuation" applications, in which patentees return to the PTO in an effort to obtain the same patents that had been denied earlier (Graham and Mowery, 2002; Lemley and Moore, 2004). To complicate matters further, continuations can sometimes result in multiple patents. Quillen, Webster and Eichman have done a careful study controlling for these variables, and they determine that the grant rate is 85 percent on the most plausible set of assumptions.

When patents are litigated, substantial uncertainty arises. Defendants in patent cases typically claim that the patent is invalid, usually based on the existence of prior art not found by the Patent and Trademark Office. However, patents are afforded a presumption of validity; to have a patent declared invalid requires “clear and convincing evidence.” Defendants also usually claim that they do not infringe the patent, even if it is valid. Of patents litigated to a final determination (appeal, trial, or summary judgment), 46 percent are held invalid (Allison and Lemley, 1998; see also Moore, 2000). To some extent, this uncertainty results because disputes that are litigated to judgment are those for which the outcome is unclear, so that the parties differ significantly in their beliefs about their prospects for winning (Priest and Klein, 1984; Cooter and Rubinfeld, 1989).⁵ But that principle cannot explain all of the uncertainty or the variation in patentee win rates over time and by court (Allison and Lemley, 1998; Chien and Lemley, 2003).

In some cases, even when patents are held valid, they are found to be not infringed or are deemed unenforceable (Moore, 2000). In some areas, particularly the process of determining the meaning of patent claims, the Federal Circuit Court of Appeals reverses district court judgments approximately one-third of the time (Chu, 2001; Moore, 2002, 2004).

A patent holder who wins an infringement suit can obtain an injunction preventing the infringing party from practicing the patent, which may force the infringing party to withdraw its products from the market. (In rare cases, the patent holder can also obtain a preliminary injunction, forcing the alleged infringer to cease using the patented technology during the patent litigation.) The infringing party might, however, be able to invent around the patent and stay in the market, albeit with higher costs or a less attractive product. A victorious patent holder also can seek damages from past infringement, either in lost profits or reasonable royalties. If the infringement is found to be “willful,” the infringing party may be forced to pay three times the actual damages.

Patents as Lottery Tickets

Once issued, a patent remains in force until 20 years after the patent application was originally filed. To keep a patent in force, the patent holder must pay maintenance fees, ranging from several hundred to a few thousand dollars, at the end of the third, seventh and eleventh years. Between 55 and 67 percent of issued U.S. patents lapse for failure to pay maintenance fees before the end of their term (Moore, 2004; Lemley, 2001), which indicates that these patents are of little value to their owners. The distribution of value of patents appears to be highly skewed, with the top 1 percent of patents more than a thousand times as valuable as the median patent (Allison, Lemley, Moore and Trunkey, 2004; Pakes, 1986; Schanker-

⁵ The magnitude of the stakes, asymmetry in the stakes and selection effects all play a role in determining observed litigation outcomes. Marco (2004) attempts to estimate these selection effects.

man and Pakes, 1986; Lanjouw and Schankerman, 1999). Many patents are virtually worthless, either because they cover technology that is not commercially important, because they are impossible to enforce effectively, or because they are very unlikely to hold up if litigated and thus cannot be asserted effectively. A small number of patents are of enormous economic significance.

Why do inventors file for many patents that turn out to have little or no value? Surely part of the reason is that patent applicants do not know which patents will be valuable and which will be worthless (Scherer, 2001; Denton and Heald, 2004). But other explanations have been offered: a failure to understand the value of patents (Rivette and Kline, 2000); the use of patents to obtain financing and boost market valuation (Lemley, 2000; Hall, Jaffe and Trajtenberg, 2005); the use of patents as signaling mechanisms (Long, 2002); and the “defensive” use of patents to deter others from suing (Hall and Ziedonis, 2001; Lemley, 2001). Even individually weak patents might have value as part of a large patent portfolio, because the portfolio can be licensed as a block or can serve to deter lawsuits (Parchomovsky and Wagner, 2004).

Many patent applications, and indeed patents themselves, are like lottery tickets. Inventors who are uncertain of the commercial significance of their ideas seek to patent many of them anyway, knowing that most of the resulting patents will turn out to be worthless, but hoping that a few will pay off big-time (Scherer, 2001). Just as people flock to stores to buy lottery tickets when the grand prize grows large, patent applicants have found ways to improve their chances of gaining patent protection in areas they consider promising. Two of the most common practices used by patentees to increase their chances of winning the patent lottery are continuations and a proliferation of closely related patents.

Patent continuations, as discussed earlier, stem from the rather remarkable rule in U.S. patent law that patent applicants are free to try again and again, without limit, to persuade the Patent and Trademark Office to grant them a patent (Lemley and Moore, 2004). They can even obtain a patent and then continue prosecution on a related application, as a hedge against the possibility that the market will change in a way that renders the first patent obsolete, or that the first patent is invalid based on prior art not cited to the PTO. Continuations are a large and growing part of patent practice, accounting for more than a quarter of all applications now filed (Graham and Mowery, 2004). In some industries, notably biotechnology and pharmaceuticals, firms typically keep a continuation application pending during the entire lifetime of the original patent (Lemley and Moore, 2004).

Patent owners also improve their chance of winning the patent lottery by filing multiple patents on closely related technologies, thereby increasing the chance that their patents will cover technology that becomes widely adopted by market participants. In a number of key industries, particularly semiconductors (Hall and Ziedonis, 2001) and computer software (Bessen and Hunt, 2004), companies file numerous patent applications on related components that are integrated into a single functional product. The result is a “patent thicket,” in which hundreds of

patents can apply to a single product (Shapiro, 2001; FTC, 2003). If the holder of a large patent portfolio asserts its patents against another company and claims that the other company is infringing dozens or even hundreds of its patents, the target company faces a very complex and costly undertaking if it chooses to fight all of those patent infringement claims in court, knowing that it has to win all or nearly all of the individual patent cases to avoid paying significant royalties or even being enjoined from selling its product (Parchomovsky and Wagner, 2004).

Both continuation practices and patent thickets can have negative consequences on other firms in the market. A competitor who designs around an issued patent—a legal activity that patent policy actively encourages (Conigliaro, Greenberg and Lemley, 2001)—cannot know whether the patentee has a continuation application waiting in the wings with claims that can be drafted to cover the design-around. Indeed, some unscrupulous patentees intentionally delay the issuance of their patents to take other firms by surprise, increasing their royalty rates once companies operating in the industry have made irreversible investments (Graham and Mowery, 2004; Lemley and Moore, 2004). Patent reformers have suggested limits on the use of continuations and courts have adopted doctrines designed to limit their abuse (FTC, 2003).

Similarly, patent thickets can have deleterious effects on both competition and innovation. One way to cut through the patent thicket is for incumbents with extensive patent portfolios to enter into broad cross-licenses (that is, exchanges of roughly symmetric patent positions) to “clear” the thicket. However, new entrants who lack large patent portfolios may be at a major disadvantage in this situation because they have no patents to trade. Without such cross-licenses, the result is inefficient “royalty stacking,” in which a manufacturer without its own patent portfolio must pay royalties to a number of separate companies.⁶ Defensive patenting is a natural, even inevitable, strategy in industries with patent thickets, but defensive patenting itself can increase the density of the thicket.

We do not mean to suggest that patent applicants invariably have little sense of the commercial significance of their inventions. To the contrary, applicants appear to have considerable private information at an early stage about the likely value of at least some of their patents. Allison, Lemley, Moore and Trunkey (2004) find that the most significant predictors of ultimate value observable to researchers are the industry, the number of prior art references, the number of claims in the patent and the time invested in prosecution of the patent. They conclude that patentees spend more time and energy along these dimensions when they believe a patent is

⁶ This is a classic instance of the “Cournot complements” problem (a form of double marginalization), which is known to lead to inefficiently high prices that can even exceed the monopoly level (Cournot, 1838; Shapiro, 2001). Heller and Eisenberg (1998) have termed a related problem the “tragedy of the anticommons.” Based on survey data, Walsh, Arora and Cohen (2003) question whether the anticommons problem has actually interfered with production in biomedical research, the area in which Heller and Eisenberg apply their theory. But among the ways they find the industry has avoided anticommons problems is to invalidate patents and to ignore them—approaches that do not deny that patents create problems in these industries.

more valuable. They quote one general counsel at a software company as saying “of the 600 patents we file a year, we pretty much know which 20 we have to have and which 580 it would be nice to have.” For the remaining majority, the lottery effect comes into play.

Reforming the System of Granting Patents

The patent system involves a *quid pro quo*. If you are the first to come up with a novel and nonobvious invention, and if you are prepared to disclose the workings of that invention to the public, in exchange you can receive exclusive rights to practice that invention for a limited time. However, when patents are improperly issued for rights that are not novel, or are “obvious,” the public suffers without justification by paying supracompetitive prices. There is widespread and growing concern that the Patent and Trademark Office issues far too many “questionable” patents that are unlikely to be found valid based on a thorough review of the sort one sees in patent litigation. Can the system can be designed to work better at reasonable cost?

Any reform of the patent system must account for the fact that patent applicants typically have superior information to the Patent and Trademark Office about likely commercial significance. There are good reasons to doubt the efficiency of a system for granting patents under which 1) patents differ greatly in their commercial significance and value; 2) patent applicants are uncertain about the value of their ideas, but have far superior information to examiners; 3) patent applicants often have superior information as well about prior art, but are under no obligation to conduct a search for the relevant prior art; 4) patent applicants can persist repeatedly through the continuation process in seeking to have certain claims accepted by patent examiners; 5) the burden of proof falls upon the PTO to explain why a patent application will *not* be granted; and 6) patent examiners are faced with a flood of applications and have little time to devote to each one. These problems are likely to be most pronounced in areas where technology is changing rapidly. Thus, the system is skewed toward the grant of patents of dubious objective validity, based on a brief, inconclusive process, which are then potentially subject to later disputes with other firms in which legal fees can easily run into millions of dollars for both sides (American Intellectual Property Law Association, 2003). Uncertainty and asymmetric information are endemic in such a patent system.

A number of scholars and policymakers have proposed reforms designed to reduce the number of improperly issued patents without causing genuine innovators to be denied patents. Two prominent recent examples come from the Federal Trade Commission (2003) and the National Academy of Sciences (2004). One common proposal is to hire more patent examiners and allow them to devote more time to the review of selected patent applications. As part of this effort, examiners could be encouraged to seek more information from patent applicants. Another common proposal is to establish a more effective opposition system in which

interested third parties could challenge the validity of an issued patent before an administrative patent board. Still another proposal is to raise the legal standard for nonobviousness, making it more difficult for applicants describing marginal improvements to obtain patent protection. We now consider some of these proposed reforms.⁷

As a starting point for thinking about patent reform, to what level of certainty should the system aspire when a patent is granted? It cannot be economically efficient to determine the validity of patents with anything approaching certainty during the application process. Indeed, Lemley (2001) has argued that the PTO is “rationally ignorant” of the actual validity of a patent, because extra resources devoted to determining the validity of a patent are largely wasted in the 95 percent of the cases in which the patent is neither litigated nor licensed for a royalty. However, others have argued that the benefits of avoiding some likely invalid patents are sufficiently great that society should spend the more money to weed out more bad patents (Gallini, 2002; Kesan and Ghosh, 2004).

Two key questions arise in considering how much effort society should put into examining patent applications. First, how effective would increased examination be at weeding out objectively bad patents without improperly denying patents for true innovations? Second, how great are the costs imposed on third parties when patents are issued that would not have been issued with more careful examination?

Some improvements in the patent examination system, like better hiring and retention practices to improve the experience and qualifications of patent examiners and altering any institutional incentives that encourage examiners to grant doubtful patents, might improve the examination process at very low cost (Merges, 1999; Thomas, 2001). However, other changes will require significant new expenditures. Devoting substantially more resources to patent examination is more likely to be efficient if those additional resources can be focused on the patents whose validity will turn out to be commercially significant. For example, following the findings of Allison, Lemley, Moore and Trunkey (2004) mentioned earlier, the patent examiners might focus greater attention on patent applications that have features correlated with greater ultimate value—like those with more claims and more prior art references. A simpler approach, which the PTO is already following to some degree, is to focus greater resources in areas of new or rapidly changing technology where the PTO has more difficulty identifying prior art, such as software and business methods. However, doing this may simply encourage patent applicants to avoid those fields (Allison and Hunter, 2004).

An alternative means of focusing attention on particular patents is to engage the incentives and information either of the party seeking the patent or parties opposed to the patent (Jaffe and Lerner, 2004). For example, patent applicants could be required to conduct a full search for prior art. New procedures could be

⁷ For a more complete analysis of the reforms proposed by the FTC (2003) and the NAS (2004), see the Summer 2004 Special Issue of the *Berkeley Technology Law Journal* devoted to these reforms, including Shapiro (2004).

established to encourage third parties to challenge a patent without entering into full-blown patent litigation. Europe and Japan already have an opposition procedure, whereby third parties can challenge issued patents in a streamlined manner. Work by Harhoff, Scherer and Vopel (2003) and others has found that the European patent opposition system is effective in identifying important patents.⁸ However, any opposition system requiring the active participation by third parties to challenge patents is subject to a free-riding problem: any third party that challenges a patent will gain only a slice of the social benefit if a patent is overturned and thus will lack sufficient incentive to challenge that patent vigorously, even if the patent is highly questionable. This general problem applies as well to the existing system of patent litigation, as discussed below (Farrell and Merges, 2004).

An alternative, more novel approach is to let patent applicants *select* either the normal, brief examination process, which would lead to a Standard Patent if the application were approved, or a more rigorous application process, which would lead to Super Patent if the application were approved. For such a system to work, the courts would have to give less weight to a Standard Patent than to a Super Patent.⁹ Patent systems based on self-selection have the attractive feature that they do not require that the Patent and Trademark Office be able to determine which technologies are most likely to be commercially significant. Indeed, one can imagine a myriad of possibilities if one thinks of the process of issuing patents in terms of designing a mechanism that can issue a variety of property rights with different levels of strength based on the process and the level of resources devoted to different patent applications.

Patent Litigation Uncertainty and Reforms

Even if the process for granting patents is improved, when a patent does enter litigation, considerable uncertainty will continue to exist about its validity and scope. The grounds for invalidating patents vary by industry, but objections related to prior art—obviousness, novelty and statutory bars—are the most common ground (Allison and Lemley, 1998). The meaning of patent claim terms—called “claim construction”—is hotly debated in virtually every patent case, and courts have found ambiguity even in such innocuous terms as “a,” “or,” “to” and “when.” Even once the meaning of the patent claims has been determined, the “doctrine of equivalents” can sometimes permit the patentee to expand its rights beyond the

⁸ Under current U.S. law, re-examinations of patents that have been granted involving third-party participation have been permitted since 1999, but remain extremely rare. According to the FTC (2003, chapter 1, p. 27) report, the *inter partes* re-examination process had been used only four times in the first three and a half years of its operation. Apparently, challengers find the opposition process unattractive, in large part because the outcome of the opposition is binding in subsequent litigation (Janis, 2000).

⁹ Doug Lichtman has suggested such an approach. Australia’s system of “petty patents” shares some features with this proposal.

literal protection of the patent (Gallini, 2002). A final source of uncertainty is the doctrine of “inequitable conduct,” in which patents may be rendered unenforceable if the patentee deceived or omitted to state information to the PTO during prosecution. While relatively few patents are held unenforceable for inequitable conduct (Moore, 2000), allegations of unenforceability are ubiquitous, and the Federal Circuit has described the growth of such claims as “an absolute plague” (*Burlington v. Dayco*, 849 F.2d 1418, 1421 [1988]).

Reform efforts focused on patent litigation have taken two very different approaches. Some reform efforts focus on reducing litigation uncertainty. For example, some reform proposals would combine a stronger process of patent examination with a stronger legal presumption of patent validity (Kesan and Banik, 2000). Similarly, the Federal Circuit has been preoccupied over the last several years with placing limits on the “doctrine of equivalents” to restrict the ability of a patent holder to expand the coverage of the patent. A number of commentators have also proposed eliminating or limiting the doctrine of “willful infringement” under which an infringer must pay enhanced damages if it intentionally infringed a patent (FTC, 2003; NAS, 2004; Lemley and Tangri, 2003; Powers and Carlson, 2001). Under current law, “willful infringement” occurs only when an infringer is aware of the patent *and* believes the patent is valid *and* believes that its conduct infringes. The law has developed a complex set of rules requiring the company to obtain an attorney’s opinion as to the proper scope and validity of the patent; reliance in good faith on that opinion will insulate a defendant from liability for willfulness. Reform proposals suggest that given the real uncertainty as to scope and validity of patents, the doctrine of willful infringement should be abolished or at least modified to put it on a more objective footing.

Other reform proposals push in the opposite direction—acknowledging that the scope and validity of patent rights are uncertain and ensuring that the law accurately reflects that uncertainty. Thus, the FTC (2003), Lemley (2001) and others have proposed legislation or judicial action to lower the burden of proof that a challenger must meet to invalidate a patent. Currently, a challenger must establish invalidity based on “clear and convincing evidence.” The FTC recommends a lower standard based on “the preponderance of the evidence.” Although it does not support the FTC’s recommendation, even the American Intellectual Property Law Association (2004) has proposed scaling back the application of the presumption of validity through judicial interpretation. Reducing the presumption of validity could also work in tandem with the Super Patents idea—Super Patents would get a strong presumption of validity, but ordinary patents would not.¹⁰

¹⁰ A few scholars have even suggested increasing uncertainty for patents whose validity is not in doubt. Ayres and Klemperer (1999) observe that the ratio of marginal profits to deadweight loss grows large as price approaches the monopoly price, and they argue that uncertain or probabilistic patent rights can therefore confer most of the private benefits of a certain right to exclude at a fraction of the social costs. They illustrate their general point using a model in which a) the patent holder would not have the right

In deciding among reform proposals, it is worth remembering that reducing litigation uncertainty is not, in and of itself, a goal in designing the patent system. Some uncertainty is an inevitable part of any system involving litigation. Furthermore, litigation over patent validity could be eliminated entirely simply by making the PTO's validity determinations final—but few would advocate such a course. In studying the uncertainty surrounding the patent system, we ultimately are interested not in that uncertainty *per se*, but rather in the effects of the patent system, and its uncertainty, on innovation, inventors, competition and consumers.

Private Incentives to Challenge Patents

If court challenges to commercially significant but questionable patents were fast and cheap, then improperly issued patents might have little market impact, because they would quickly be challenged and overturned. Unfortunately, the patent litigation process does not work in anything approaching this idealized fashion.

The main problem with the litigation system can be demonstrated with an example. Suppose that widgets are supplied in a competitive market consisting of ten identical firms, each with constant marginal cost of \$40 per unit. For simplicity, suppose that demand for widgets is linear, given by $P = 100 - Q$, where P is price and Q is quantity. The resulting competitive price of widgets is \$40, and 60 widgets are sold. Now suppose that a new method of producing widgets is developed which lowers the production cost to \$30 per unit. If this method is freely available to all producers, the price of widgets will fall to \$30 per unit, and quantity will rise to 70. Suppose, however, that a patent is issued covering this new method of production. Again for simplicity, suppose that the patent is held by a firm that does not produce widgets. The owner of this patent selects a royalty rate, R , in dollars per widget, at which it will license its patent to all widget manufacturers.

Begin with the case in which the patent is unquestionably valid. In this case, widget manufacturers have a choice to make: continue producing at \$40 per unit or take a license, produce at \$30 per unit using the new and superior method of production, but pay royalties. From the perspective of the patent holder, setting the optimal royalty rate is a problem of monopoly pricing. No one will pay a royalty greater than \$10 per unit, since that is the cost savings associated with the patented process. Indeed, the monopoly outcome is for the patent holder to choose a royalty

to stop a rival from infringing, but rather could only seek compensatory damages after the patent expired, and b) even these damages would only be awarded with some probability. While we agree that conventional patent policy is “inefficient at the margin” in the sense they describe, it is not clear to us how their proposal could plausibly be made operational.

$R^* = 10$.¹¹ The patent holder receives royalty payments of \$10 per unit for 60 units, appropriating the entire cost savings associated with its invention. There is a deadweight loss of \$50 resulting from the fact that only 60 units are produced, not 70, which would be socially optimal given the actual (social) cost of production of \$30 per widget.¹²

Now change the story to reflect probabilistic real-world patents. Suppose that the ten widget producers are aghast that this patent was issued and insist that the patented method was obvious to someone skilled in art at the time of the patent application. Indeed, we may imagine that, following the publication of some basic research in the public domain, the widget producers soon learned how to apply those research findings to their production methods and thus lower their costs by \$10 per unit. Perhaps they even made significant, technology-specific capital investments. Meanwhile, unknown to them, someone outside the industry had filed for a patent covering this technology and convinced the PTO that its application met the novelty and nonobviousness tests. For simplicity, let us suppose that the patent at issue is highly questionable: there is only a 20 percent chance that the patent would be found valid if tested in court. However, litigation involves some cost, \$C. Will any of the widget manufacturers challenge the patent?

If any individual widget maker considers challenging the patent, it must consider two outcomes. If the patent is upheld, then the firm has spent \$C and gained nothing. If the patent is overturned, then all firms will be relieved of any royalty obligations, and the price of widgets will fall to \$30.¹³ Invalidating the patent is a public good that benefits consumers of widgets, but not any one widget manufacturer or even widget manufacturers collectively. Thus, no individual widget maker can recover the \$C litigation costs. Accepting the \$10 royalty is a dominant strategy for each widget manufacturer in this setting. In the end, consumers end up paying \$10 per widget to the patent holder even if the patent should never have been issued. Furthermore, the prospect of the prize of \$600 in royalties to the patent holder will encourage rent-seeking behavior by patent applicants.

What is driving this striking result that even a weak patent can command royalties approaching those of an ironclad patent covering the same claims? The key insight is that invalidating a patent generates significant positive externalities, and activities that generate positive externalities are undersupplied. There are very strong reasons to believe that challenges to patents are undersupplied (Gilbert, 2004; Farrell and Merges, 2004). In practice, this means that companies accused of infringing will tend to settle patent disputes—for example, by paying royalties—

¹¹ For a royalty rate R between zero and ten, the price of widgets will be $30 + R$, and the quantity produced will be $70 - R$. The patent owner's licensing revenues will be $R(70 - R)$. This expression is strictly increasing in R for R between zero and ten.

¹² The \$50 represents ten units not produced, with an average foregone social surplus of \$5 per unit. Each unit would have cost \$30 to produce, and their average value to consumers would have been \$35.

¹³ The key case is the Supreme Court's decision in *Blonder-Tongue Lab v. University of Illinois Found* (402 U.S. 313 [1971]), under which an alleged infringer can prevent an infringement suit if the patent claim asserted against it has been declared invalid in another case.

rather than litigating. It may also lead companies to pool their patents inefficiently and share royalties rather than engage in patent litigation (Gilbert, 2004; Choi, 2003). In a more general setting, Farrell and Shapiro (2005) show that the royalties commanded by the owner of a probabilistic patent can easily be disproportionate to the strength of the patent and are highest when a single licensee's profits are very sensitive to its own costs but not sensitive to the level of industry-wide costs.

We do not mean to suggest that our simple example provides anything approaching a complete analysis of this problem, nor that the royalties commanded by a patent are generally unrelated to the strength of that patent. In the example just given, the widget manufacturers had no incentive at all to challenge patents because their position is symmetrical and competition is perfect. In the real world, participants in most industries have rents stemming from imperfections in competition, specific capital investments, oligopoly, product differentiation and brand value and nonconstant marginal cost curves. As a result, most accused infringers will have *some* incentive to challenge the validity of a patent, but that incentive will be suboptimal. Whether an accused infringer has *sufficient* incentive to mount an effective challenge will depend on a number of factors: how significant are those rents; how readily they will be dissipated to competitors; how much money is at stake based on past infringement; and what is the relationship between litigation expenditures and success at trial? A company sued for \$1 billion in royalties will likely have the incentive to pay \$5 or \$10 million in legal fees even if competitors will also benefit substantially; a company sued for \$10 million in royalties may well not have strong incentives to defend the suit. In short, the simple example just presented is a polar case. But more general analyses show that serious problems arise in relying on private parties to challenge questionable patents.

Just starting from the example sketched above, a number of questions spring to mind. How would the analysis change if the widget manufacturers produced differentiated products? How would the analysis change if the widget manufacturers had made specific capital investments and produced subject to increasing marginal cost, at least in the short run? How would the outcome change if existing manufacturers could coordinate their decisions to challenge the patent (while still being prohibited from colluding on price)? How would the analysis change if only a few widget makers existed and they acted as oligopolists? Can downstream consumers, the ultimate beneficiaries of a successful patent challenge in this model, effectively band together to fight the patent? On the other hand, could the patent holder fight back by credibly threatening to charge a higher royalty to anyone challenging its patent than to those who agree to pay royalties without a fight? Farrell and Merges (2004) offer insightful further discussion of these issues. They emphasize two basic reasons why individual firms accused of patent infringement have suboptimal incentives to challenge the patents asserted against them: 1) the public good problem—the fact that rivals to the allegedly infringing firm will benefit from a finding that the patent is invalid (or that its claims should be read narrowly); and 2) the pass-through problem—the fact that higher uniform royalty costs are passed through in the form of higher prices, thus muting the incentives

of alleged infringers to avoid paying such uniform royalties.¹⁴ Farrell and Shapiro (2005) ask more generally how holders of weak patents will structure their licensing agreements to induce licensees to accept licenses rather than to challenge their patents.

Since invalidating a patent provides a public good, typically to the benefit of competitors and consumers, one can naturally consider policies to overcome this public-good problem. One standard approach for dealing with public goods is to reward or subsidize those who contribute to the public good. For example, one might subsidize those who successfully challenge patents by instituting a bounty system (Thomas, 2001; Miller, 2004). An alternative reward would be to give certain exclusive rights to the party who successfully challenges a patent. The Hatch-Waxman Act, which grants a limited period of exclusivity to the first generic supplier to challenge a pharmaceutical patent, has this flavor. These incentives are counterbalanced by the fact that a company that initiates a post-grant opposition signals to the patent holder that it may be infringing the patent. In the words of industry participants, a firm that initiates a post-grant opposition effectively “paints a big target on its back.”

A second standard approach is for the government to supply the public good. The government can and does challenge some issued patents when the PTO re-examines a patent, perhaps in response to third party complaints or information. However, such government-led re-examinations are currently rare. One way around this problem would be to empower a government agency to challenge patents based on information provided by interested industry participants, even ones who remain anonymous. The Federal Trade Commission, with its consumer-oriented mission, might be suitable for this task, perhaps working in conjunction with other agencies that could contribute their technical expertise in different areas of science and technology. Another approach is to encourage public interest organizations to challenge suspect patents. Two such organizations, the Electronic Frontier Foundation and the Public Patent Foundation, have begun to file administrative challenges to patents.

A final approach is to impose restrictions on the agreements that litigants can reach to settle patent cases, in order to prevent agreements that harm competition. We discuss such restrictions in the next section.

¹⁴ Farrell and Merges (2004) also assert that the outcome of patent litigation tends to be tilted toward the party spending more on litigation, which tends to be the party with the most at stake. Therefore, if the patent holder has much more at stake than does any individual alleged infringer, even if litigation occurs, the outcome may be tilted in favor of the patent holder. While this may well be true as a general matter, we strongly suspect that expenditures on litigation are subject to diminishing returns, so such differences may be of little significance in high-stakes cases. A firm spending \$1 million litigating a patent case will likely do much better than a firm that spends only \$100,000, but it does not follow that a firm spending \$20 million will do much better than a firm spending \$15 million.

Antitrust Limits on Patent Settlements

Of the 1.5 percent of all patents that are litigated, some 95 percent of the cases end in settlements rather than verdicts. This figure does not count all of the settlements that happen without suit ever being filed. The prevalence of actual and potential competitors entering into patent settlement agreements that restrict competition raises important questions about possible antitrust violations.

Some antitrust limits on the settlements of patent disputes between rivals are unquestionably needed. Consider an incumbent monopolist who faces the threat of entry from a single potential entrant. The monopolist owns a patent that it is asserting against the potential entrant. For simplicity, assume that both parties agree that if fully litigated, the patent will be found valid with probability P . Suppose that the patent at issue is valuable only in this market, and that no other potential entrants exist, so no public-goods problem arises with invalidating this patent. However, consumers are affected by the presence or absence of competition, so externalities remain associated with the decision to litigate or settle and the terms on which a settlement occurs. There is no reason to assume that bargaining between the monopolist and the potential entrant to maximize their joint profits will lead to a socially optimal settlement.

Indeed, the incumbent monopolist and the potential entrant will quite probably achieve an anticompetitive settlement, at least in the absence of antitrust rules limiting the manner in which they can resolve their dispute. As long as monopoly profits are greater than joint duopoly profits, the monopolist and the entrant will have an incentive to negotiate in a way that leads to the monopoly level of output and the monopoly price. In comparison with litigation, such a settlement would deprive consumers of the competition that would arise if the patent were declared invalid, which would occur with probability $1 - P$. For a relatively weak patent, consumers can be significantly harmed by an agreement between the incumbent monopolist and the potential entrant that maximizes their joint profits. In the limiting case as P approaches zero, a weak patent can be used as a fig leaf to cover an agreement not to compete.

The expected joint profits of the incumbent monopolist and potential entrant are higher from a clever settlement agreement than from litigation. Expected profits from settlement depend upon the terms of the settlement. But consider an easy way for the parties to settle and achieve full monopoly profits: the incumbent can pay the potential entrant not to enter the market.¹⁵ Intuitively, settlement leads to higher joint profits for two reasons: settlement eliminates the chance that

¹⁵ How big a payment is needed? By litigating the entrant, E can earn expected profits of $(1 - P) \times \pi_E - C_E$, where π_E represents E 's profits and C_E represents E 's litigation costs. The incumbent M can induce E to agree not to enter the market by paying E an amount $F > (1 - P) \times \pi_E - C_E$. Is there a mutually agreeable level of F ? Yes. Including the payment of F , M earns $\pi_M - F$, where π_M represents the incumbent's profits if there is no entry. In contrast, by litigating, M would earn expected profits of $P \times \pi_M + (1 - P) \times \pi_I - C_M$, where π_I represents the incumbent's profits if entry occurs. Settling is superior to litigating for M if and only if $\pi_M - F > P \times \pi_M + (1 - P) \times \pi_I - C_M$, which can be written as $F < (\pi_M$

profit-dissipating competition will break out if the patent is proven invalid, and it avoids litigation costs.¹⁶ A particularly corrosive form of settlement occurs when the litigants collude to ask the court to vacate a decision it has already rendered, often one holding a patent invalid or construing its scope narrowly. A published court decision on the validity or scope of a patent is a public good, and courts should refuse to vacate their opinions upon settlement. Nonetheless, judges sometimes do vacate earlier decisions at the request of the litigating parties.

Of course, it is blatantly illegal for a monopolist to pay its sole potential competitor to stay off the market. Nor could such a payment be hidden in the form of an acquisition: under U.S. antitrust law, a monopolist is not allowed to acquire its sole potential entrant. But is this same payment anticompetitive in the context of a patent settlement? Does the answer depend upon the strength of the patent?

The courts have been grappling with this issue, and more generally the antitrust limits on patent settlements, in a series of cases over the past five years (Hovenkamp, Janis and Lemley, 2003, 2004b; Cotter, 2003). Many of these cases have arisen from settlements between incumbent pharmaceutical manufacturers and potential generic competitors who they have accused of patent infringement, largely because the Hatch-Waxman legislation regulates entry into the market and therefore gives a pharmaceutical patentee who settles with a generic some power to keep out all competitors, not just one (FTC, 2002; Morse, 2002).

The courts have devoted most of their attention to settlements involving payments from incumbents to would-be generic suppliers, known as “reverse payments” because they flow from the patent holder to the challenger, in contrast to conventional licensing payments that challengers make to patent holders. Courts have come to different conclusions on the legality of such reverse or “exclusion payments.” For example, Abbott, the maker of Hytrin, a very successful drug used to treat hypertension and enlarged prostate, was faced with potential generic entry by Geneva, which Abbott accused of infringing its patent. Under their 1998 settlement, Abbott agreed to pay Geneva \$4.5 million per month for some period of time in exchange for which Geneva agreed not to enter the market. The Eleventh Circuit found that this agreement was not *per se* illegal and instructed the

$-\pi_I) \times (1 - P) + C_M$. There are mutually attractive levels of F if and only if $(1 - P)^* \pi_E - C_E < (\pi_M - \pi_I) \times (1 - P) + C_M$. Rearranging, this is equivalent to

$$(\pi_M - \pi_I - \pi_E) \times (1 - P) + C_M + C_E > 0.$$

Assuming $\pi_M > \pi_I + \pi_E$, i.e., that competition dissipates total profits, this inequality must be satisfied.
¹⁶ These dangers are not limited to the case of a single potential entrant. With multiple potential entrants, the firms still benefit as a group if they can agree to licensing terms that support the monopoly outcome. However, the presence of multiple potential licensees requires much more complex multi-party negotiations. We do not mean to suggest that the monopoly outcome necessarily can be achieved in all cases. Among other things, it requires significant barriers to entry in order to work. See Farrell and Shapiro (2005) for further analysis along these lines.

District Court to reconsider the case.¹⁷ A variant of this pattern arose in a case involving Schering-Plough, the maker of the prescription drug K-Dur 20, which is used to treat low potassium. Schering-Plough entered into an agreement with Upsher-Smith, a potential generic competitor, which involved a payment from Schering-Plough to Upsher-Smith and an agreement by Upsher-Smith not to enter the market before a specified date. The Federal Trade Commission has found this agreement to be anticompetitive and held that agreements of this sort are presumptively unlawful. The Eleventh Circuit reversed. The courts have yet to establish a clear and uniform approach to these cases.¹⁸

Previous work by the authors and others has argued that patent settlements involving payments in excess of avoided litigation costs by incumbent patent holders to potential entrants accused of infringing should be presumed to be anticompetitive (Hovenkamp, Janis and Lemley, 2003; Shapiro, 2003a, 2003b; Cotter, 2003, 2004; Leffler and Leffler, 2002, 2003). The likely effect of such payments is to delay entry, either in comparison with the outcome of litigation over the patent or in comparison with a settlement not involving these payments.¹⁹ Neither competition nor innovation is promoted by allowing the owner of a weak patent to pay would-be challengers to refrain entirely from competing during the lifetime of the patent.

Defenders of such agreements, and some courts, have argued that the antitrust laws should not be used to weaken patent protection, and that the patent at issue might well not be invalid (Schildkraut, 2004; Crane, 2004). We agree that a patent holder does not violate the antitrust laws by excluding rivals who have been proven to infringe a valid patent; the essence of a patent is a right to exclude others from practicing the patented invention. But a patent does not give its owner the right to exclude rivals who are *allegedly* infringing, at least not without a court order. Payments from patent holders to *alleged* infringers in exchange for their agreement to stay off the market therefore go beyond the patent grant and exclude *allegedly* infringing competition, to the detriment of consumers (Shapiro, 2003b). If a

¹⁷ *Valley Drug Co. v. Geneva Pharmaceuticals, Inc.*, 344 F. 3rd 1294 (11th Cir. 2003). On remand, the district court once again found the agreement illegal. Shapiro has served as an expert witness for Kaiser, a purchaser of Hytrin who claims to be injured by Abbott's agreement with Geneva.

¹⁸ The decision of the Federal Trade Commission in this case, *In the Matter of Schering-Plough Corporation, et al.*, is available at (<http://www.ftc.gov>). In contrast to the Abbott-Geneva agreement just noted, the Sixth Circuit found a similar agreement between Hoescht Marion Roussel and Andrx regarding the drug Cardizem CD to be *per se* illegal (*In re Cardizem Antitrust Litigation*, 3322 F. 3rd 896, Sixth Circuit, 2003). In yet another case, the District Court ruled that a similar agreement involving Cipro was not *per se* illegal because the strength of the patent must be considered as part of the antitrust analysis (*In re Ciprofloxacin Hydrochloride Antitrust Litigation*, 261 F. Supp. 2nd 188, E.D.N.Y. 2003). This court later found the Cipro agreement to be lawful.

¹⁹ Willig and Bigelow (2004) argue that in cases involving negotiated entry dates, such reverse payments can be procompetitive. They thus oppose a *per se* rule prohibiting such payments. Under the standard we favor, such payments would be presumptively anticompetitive, giving the settling parties the chance to demonstrate in a given case that justifications such as those offered by Willig and Bigelow—based on risk aversion, imperfect capital markets and asymmetric information—apply with sufficient force to overcome the presumption.

patent holder believes that it properly has the right to exclude an alleged infringer, the patent holder can seek a preliminary injunction forcing the alleged infringer from the market. Such preliminary injunctions may be granted if the patent holder is likely to prevail and would suffer irreparable harm from the allegedly infringing competition, perhaps because the alleged infringer would not be able fully to compensate the patent owner for damages if the patent is later found to be valid and infringed. Indeed, in the pharmaceutical context, in which every reverse payment case so far has arisen, the law provides for an *automatic* preliminary injunction for 30 months; reverse payments are an effort to extend the period of exclusion beyond that point without having to litigate the patent. Importantly, courts will deny a preliminary injunction if serious questions exist about the validity of the patent, so the patentee's decision to settle with a reverse payment rather than to seek an injunction provides some evidence about the perceived strength of the patent. In addition, we show in our previous articles that one often can infer a certain degree of patent weakness from the fact that the patent owner is paying to avoid the risk that its patent will be found invalid, especially if the payment is large. Therefore, large reverse payments are inconsistent with a claim by the patent holder that its patent very likely would be found valid if litigated.

Patent settlements take many forms and can raise a variety of antitrust issues even when they do not involve reverse payments from patent holders to alleged infringers to stay off the market. Virtually every licensing agreement can be seen as the settlement of a potential patent dispute. In other cases, patent disputes are settled through mergers or acquisitions or by forming patent pools or engaging in cross-licensing. Some pharmaceutical companies have settled with would-be generic suppliers by negotiating a date at which the generic supplier could enter the market; Hovenkamp, Janis and Lemley (2003) have argued that such agreements are reasonable so long as they do not include a reverse payment. Patent applicants also settle interference disputes in which each claims to be the first to have come up with an invention; such settlements will typically involve a payment from one side to the other because the side conceding priority is giving up potential ownership of the patent (Hovenkamp, Janis and Lemley, 2004a).

Most settlements are quite reasonable in competitive terms, and thus we face the complex question of identifying those few settlements that are anticompetitive. Shapiro (2003a) proposed a framework for establishing the antitrust limits on patent settlements based on the following principle: the settlement cannot lead to lower expected consumer surplus than would arise from ongoing litigation. This rule is similar in character to existing antitrust rules relating to mergers and to licensing agreements, as reflected in the *Horizontal Merger Guidelines* and the *Anti-trust Guidelines for the Licensing of Intellectual Property* issued by the Federal Trade Commission and the U.S. Department of Justice. The rule proposed and analyzed by Shapiro also respects the rights of patent holders, while preventing companies from using the cover of patent settlements to engage in cartel-like agreements. For better or worse, applying this rule typically requires some assessment of the strength of the relevant patents, either directly or by inference. In the "reverse payment"

cases described above, the presence of a substantial payment from patentee to accused infringer may well imply that the patentee paid for a reduction in competition in comparison with ongoing litigation. In contrast, this same framework implies that simple licensing agreements cannot be presumed anticompetitive: royalties will reflect the underlying strength of the patent, since no licensee will burden itself with high royalties unless the patent is indeed likely to be found valid and infringed. Shapiro (2003a) studies a variety of other, more complex types of settlements, including mergers and patent pools as well as negotiated entry dates.

Conclusion

The patent system does not grant an absolute right to inventors to exclude others from practicing their inventions, as many economic models assume. Rather, the patent system gives the patent holder a right to *try* to exclude others by asserting its patent against them in court. The actual scope of a patent right, and even whether the right will withstand litigation at all, are uncertain and contingent questions. This uncertainty is not an accident or mistake. Rather, it is an inherent part of our patent system, an accommodation to the hundreds of thousands of patent applications filed each year, the inability of third parties to participate effectively in determining whether a patent should issue, and the fact that for the vast majority of issued patents, scope and validity are of little or no commercial significance. Modeling patents as probabilistic rights requires us to rethink reform of the patent granting process, patent opposition procedures, our approach to patent litigation, the efficacy of litigation as a means of invalidating patents that were improperly issued and antitrust policy toward the settlement of patent lawsuits.

■ We thank Rochelle Dreyfuss, Joe Farrell, Richard Gilbert, Rose Hagan, James Hines, Paul Klemperer, Josh Lerner, Kevin Outterson, Timothy Taylor and Michael Waldman for their input, including comments on an earlier draft.

References

- Allison, John R. and Starling Hunter. 2004. "On Experimental Patent Reform and the Failure to Fix a Non-Problem." Working paper.
- Allison, John R. and Mark A. Lemley. 1998. "Empirical Analysis of the Validity of Litigated Patents." *American Intellectual Property Law Association Quarterly Journal*. 26:3, pp. 185–275.
- Allison, John R. and Mark A. Lemley. 2000. "Who's Patenting What? An Empirical Exploration of Patent Prosecution." *Vanderbilt Law Review*. 53:6, pp. 2099–174.
- Allison, John R., Mark A. Lemley, Kimberly A. Moore and R. Derek Trunkey. 2004. "Valuable Patents." *Georgetown Law Journal*. 92:3, pp. 435–80.
- American Intellectual Property Law Association. 2003. *Report of Annual Economic Sur-*

vey. Arlington, Va.: American Intellectual Property Law Association.

American Intellectual Property Law Association. 2004. Response to the October 2003 FTC Report; Available at (<http://www.aipla.org>).

Ayres, Ian and Paul Klemperer. 1999. "Limiting Patentees' Market Power Without Reducing Innovation Incentives: The Perverse Benefits of Uncertainty and Non-Injunctive Remedies." *Michigan Law Review*. 97:4, pp. 985–1033.

Bessen, James and Robert M. Hunt. 2004. "An Empirical Look at Software Patents." Working Paper No. 03-17; Available at (<http://www.researchoninnovation.org/swpat.pdf>).

Chien, Colleen V. and Mark A. Lemley. 2003. "Are the U.S. Patent Priority Rules Really Necessary?" *Hastings Law Journal*. 54:5, pp. 1299–333.

Choi, Jay Pil. 2003. "Patent Pools and Cross-Licensing in the Shadow of Patent Litigation." CESifo Working Paper No. 1070.

Chu, Christian A. 2001. "Empirical Analysis of the Federal Circuit's Claim Construction Trends." *Berkeley Technology Law Journal*. 16:3, pp. 1075–164.

Conigliaro, Matthew, Andrew Greenberg and Mark A. Lemley. 2001. "Foreseeability in Patent Law." *Berkeley Technology Law Journal*. 16:3, pp. 1045–074.

Cooter, Robert and Daniel Rubinfeld. 1989. "Economic Analysis of Legal Disputes and their Resolution." *Journal of Economic Literature*. 27:4, pp. 1067–097.

Cotter, Thomas F. 2003. "Refining the 'Pre-sumptive Illegality' Approach to Settlements of Patent Disputes Involving Reverse Payments." *Minnesota Law Review*. 87:6, pp. 1789–816.

Cotter, Thomas F. 2004. "Antitrust Implications of Patent Settlements Involving Reverse Payments: Defending a Rebuttable Presumption of Illegality in Light of Some Recent Scholarship." *Antitrust Law Journal*. 71:3, pp. 1069–097.

Cournot, A. A. 1838. *Researches into the Mathematical Principles of the Theory of Wealth*. English Edition, New York: Kelley.

Crane, Daniel A. 2004. "Ease Over Accuracy in Assessing Patent Settlements." *Minnesota Law Review*. 88:3, pp. 698–711.

Denton, F. Russell and Paul J. Heald. 2004. "Random Walks, Non-Cooperative Games and the Complex Mathematics of Patent Pricing." *Rutgers Law Review*. 55:4, pp. 1175–288.

Dreyfuss, Rochelle Cooper. 1989. "The Federal Circuit: A Case Study in Specialized Courts." *New York University Law Review*. 64:1, pp. 1–74.

Farrell, Joseph and Robert P. Merges. 2004. "Incentives to Challenge and Defend Patents: Why Litigation Won't Reliably Fix Patent Office

Errors and Why Administrative Patent Review Might Help." *Berkeley Technology Law Journal*. 19:3, pp. 943–70.

Farrell, Joseph and Carl Shapiro. 2005. "How Strong Are Weak Patents?" Working paper, University of California at Berkeley.

Federal Trade Commission (FTC). 2002. "Generic Drug Entry Prior to Patent Expiration: An FTC Study." Available at (<http://www.ftc.gov/os/2002/07/genericdrugstudy.pdf>).

Federal Trade Commission (FTC). 2003. "To Promote Innovation: A Proper Balance of Competition and Patent Law and Policy." Available at (<http://www.ftc.gov/opa/2003/10/cpreport.htm>).

Gallini, Nancy T. 2002. "The Economics of Patents: Lessons From Recent U.S. Patent Reform." *Journal of Economic Perspectives*. 16:2, pp. 131–54.

Gilbert, Richard J. 2004. "Antitrust for Patent Pools: A Century of Policy Evolution." *Stanford Technology Law Review*. April; Available at (http://stlr.stanford.edu/STLR/Core_Page/).

Graham, Stuart and David Mowery. 2004. "Submarines in Software? Continuations in US Software Patenting in the 1980s and 1990s." *Economics of Innovation and New Technology*. 13:5, pp. 443–56.

Hall, Bronwyn H. and Rosemarie Ziedonis. 2001. "The Patent Paradox Revisited: An Empirical Study of Patenting in the U.S. Semiconductor Industry, 1979–1995." *RAND Journal of Economics*. 32:1, pp. 101–28.

Hall, Bronwyn H., Adam B. Jaffe and Manuel Trajtenberg. 2005. "Market Value and Patent Citations." *RAND Journal of Economics*. 36:1, forthcoming.

Harhoff, Dietmar and Markus Reitzig. 2004. "Determinants of Opposition Against EPO Patent Grants—The Case of Biotechnology and Pharmaceuticals." *International Journal of Industrial Organization*. 22:4, pp. 443–80.

Harhoff, Dietmar, Frederic M. Scherer and Katrin Vopel. 2003. "Citations, Family Size, Opposition and the Value of Patent Rights." *Research Policy*. 328, pp. 1343–363.

Heller, Michael and Rebecca Eisenberg. 1998. "Can Patents Deter Innovation? The Anticommons in Biomedical Research." *Science*. 280:5364, pp. 698–701.

Hovenkamp, Herbert, Mark D. Janis and Mark A. Lemley. 2003. "Anticompetitive Settlement of Intellectual Property Disputes." *Minnesota Law Review*. 87:6, pp. 1719–766.

Hovenkamp, Herbert, Mark D. Janis and Mark A. Lemley. 2004a. *IP and Antitrust: An Analysis of*

Antitrust Principles Applied to Intellectual Property Law. New York: Aspen Publishers.

Hovenkamp, Herbert, Mark D. Janis and Mark A. Lemley. 2004b. "Balancing Ease and Accuracy in Assessing Pharmaceutical Exclusion Payments." *Minnesota Law Review*. 88:3, pp. 712–21.

Jaffe, Adam B. and Josh Lerner. 2004. *Innovation and Its Discontents: How Our Broken Patent System is Endangering Innovation and Progress, and What to Do About It*. Princeton, N.J.: Princeton University Press.

Kaplow, Louis. 1984. "The Patent-Antitrust Intersection: A Reappraisal." *Harvard Law Review*. June, 97, pp. 1813–889.

Kesan, Jay P. and Marc Banik. 2000. "Patents as Incomplete Contracts: Aligning Incentives for R&D Investment With Incentives to Disclose Prior Art." *Washington University Journal of Law and Policy*. 2:23, pp. 23–54.

Kesan, Jay P. and Shubha Ghosh. 2004. "What Do Patents Purchase? In Search of Optimal Ignorance in the Patent Office." *Houston Law Review*. 40, pp. 1219–264.

Kortum, Samuel and Joshua Lerner. 1999. "Stronger Protection or Technological Revolution: What is Behind the Recent Surge in Patenting?" *Research Policy*. 28:1, pp. 1–22.

Lanjouw, Jean O. and Mark Schankermann. 2001. "Characteristics of Patent Litigation: A Window on Competition." *RAND Journal of Economics*. 32:1, pp. 129–51.

Lanjouw, Jean O. and Mark Schankerman. 2004. "Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators." *Economic Journal*. 114:495, pp. 441–65.

Lanjouw, Jean O., Ariel Pakes and Jonathan Putnam. 1998. "How to Count Patents and Value Intellectual Property: Uses of Patent Renewal and Application Data." *Journal of Industrial Economics*. 44:4, pp. 405–33.

Leffler, Keith and Cristofer Leffler. 2002. "Want to Pay a Competitor to Exit the Market? Settle a Patent Infringement Case." *Economics Committee Newsletter*. 2:1, pp. 26–35.

Leffler, Keith and Cristofer Leffler. 2003. "The Probabilistic Nature of Patent Rights: In Response to Kevin McDonald." *Antitrust*. Summer, pp. 77–83.

Lemley, Mark A. 2000. "Reconceiving Patents in the Age of Venture Capital." *Journal of Small and Emerging Business Law*. 4:1, pp. 137–48.

Lemley, Mark A. 2001. "Rational Ignorance at the Patent Office." *Northwestern University Law Review*. 95:4, pp. 1497–532.

Lemley, Mark A. and Kimberly A. Moore. 2004. "Ending Abuse of Patent Continuations." *Boston University Law Review*. 84:1, pp. 63–123.

Lemley, Mark A. and Ragesh K. Tangri. 2003. "Ending Patent Law's Willfulness Game." *Berkeley Technology Law Journal*. 18:4, pp. 1085–126.

Lerner, Joshua. 1995. "Patenting in the Shadow of Competitors." *Journal of Law and Economics*. 38:2, pp. 463–95.

Lerner, Joshua. 2002. "Where Does *State Street* Lead? A First Look at Finance Patents, 1971–2000." *Journal of Finance*. 57:2, pp. 901–30.

Long, Clarisa. 2002. "Patent Signals." *University of Chicago Law Review*. 69:2, pp. 625–79.

Marco, Alan C. 2004. "The Selection Effects (and Lack Thereof) in Patent Litigation: Evidence from Trials." *Topics in Economic Analysis and Policy*. 4:1, article 21.

Merges, Robert P. 1999. "As Many As Six Impossible Patents Before Breakfast: Property Rights for Business Concepts and Patent System Reform." *Berkeley Technology Law Journal*. 14:2, pp. 577–616.

Miller, Joseph Scott. 2004. "Building a Better Bounty: Litigation-Stage Rewards for Defeating Patents." *Berkeley Technology Law Journal*. 19:2, pp. 667–739.

Moore, Kimberly A. 2000. "Judges, Juries and Patent Cases—An Empirical Peek Inside the Black Box." *Michigan Law Review*. 99:2, pp. 365–409.

Moore, Kimberly A. 2002. "Are District Court Judges Equipped to Handle Patent Cases?" *Harvard Journal of Law and Technology*. 15:1, pp. 1–39.

Moore, Kimberly A. 2004. "Worthless Patents." Available at (<http://mason.gmu.edu/~kamoore/>).

Morse, M. Howard. 2002. "Settlement of Intellectual Property Disputes in the Pharmaceutical and Medical Device Industries: Antitrust Rules." *George Mason Law Review*. 10:3, pp. 359–405.

National Academies of Science (NAS). 2003. *Patents in the Knowledge-Based Economy*. Wesley M. Cohen and Stephen A. Merrill, eds. Washington, D.C.: National Academies Press.

National Academies of Science. 2004. "A Patent System for the 21st Century." Available at (<http://www.nap.edu/books/0309089107/html>).

Nordhaus, William. 1969. *Invention, Growth and Welfare: A Theoretical Treatment of Technological Change*. Cambridge, Mass.: MIT Press.

Pakes, Ariel. 1986. "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks." *Econometrica*. 54:4, pp. 755–784.

Parchomovsky, Gideon and R. Polk Wagner. 2004. "Patent Portfolios." Working paper, University of Pennsylvania Law School.

Powers, Matthew D. and Stephen C. Carlson.

2001. "The Evolution and Impact of the Doctrine of Willful Patent Infringement." *Syracuse Law Review*. 51:1, pp. 53–113.

Priest, George L. and Benjamin Klein. 1984. "The Selection of Disputes for Litigation." *Journal of Legal Studies*. 13:1, pp. 1–55.

Quillen, Cecil D., Jr., Ogden H. Webster and Richard Eichmann. 2003. "Continuing Patent Applications and Performance of the U.S. Patent and Trademark Office—Extended." *Federal Circuit Bar Journal*. 12:1, pp. 35–55.

Reinganum, Jennifer. 1989. "The Timing of Innovation: Research, Development and Diffusion," in *The Handbook of Industrial Organization*. Robert Willig and Richard Schmalensee, eds. Amsterdam: North-Holland, pp. 849–908.

Rivette, Kevin and David Kline. 2000. *Rembrandts in the Attic: Unlocking the Hidden Value of Patents*. Boston: Harvard Business School Press.

Schankerman, Mark and Ariel Pakes. 1986. "Estimates of the Value of Patent Rights in European Countries During the Post-1950 Period." *Economic Journal*. 96:127, pp. 1052–076.

Schechter, Roger E. and John R. Thomas. 2003. *Intellectual Property: The Law of Copyrights, Patents and Trademarks*. St. Paul, Minnesota: Thomson West.

Scherer, F. M. 2001. "The Innovation Lottery," in *Expanding the Boundaries of Intellectual Property*. Rochelle Dreyfuss et al., eds. Oxford: Oxford University Press, pp. 3–21.

Schildkraut, Marc G. 2004. "Patent-Splitting Settlements and the Reverse Payment Fallacy." *Antitrust Law Journal*. 71:3, pp. 1033–068.

Segal, Ilya. 1999. "Contracting with Externalities." *Quarterly Journal of Economics*. 104:2, pp. 337–88.

Shapiro, Carl. 2001. "Navigating the Patent Thicket: Cross-Licenses, Patent Pools, and Standard-Setting." in *NBER Innovation Policy and the Economy*. Adam Jaffe, Joshua Lerner and Scott Stern, eds. Cambridge, Mass.: MIT Press, pp. 119–50.

Shapiro, Carl. 2003a. "Antitrust Limits to Patent Settlements." *RAND Journal of Economics*. 34:2, pp. 391–411.

Shapiro, Carl. 2003b. "Antitrust Analysis of Patent Settlements Between Rivals." *Antitrust*. Summer, pp. 70–77.

Shapiro, Carl. 2004. "Economic Critique of the FTC's Proposals for Patent Reform." *Berkeley Technology Law Journal*. 19:23, pp. 943–70.

Thomas, John R. 2001. "Collusion and Collective Action in the Patent System: A Proposal for Patent Bounties." *University of Illinois Law Review*. 2001:1, pp. 305–53.

U.S. Patent and Trademark Office. 2003. "Annual Report for 2003." Available at <http://www.uspto.gov/web/offices/com/annual/2003/2003annualreport.pdf>.

U.S. Patent and Trademark Office. 2003. "The 21st Century Strategic Plan 4." February 3; Available at <http://www.uspto.gov/web/offices/com/strat21/index.htm>.

Walsh, John P., Ashish Arora and Wesley M. Cohen. 2003. "Effects of Research Tool Patents and Licensing on Biomedical Innovation," in *Patents in the Knowledge-Based Economy*. Wesley M. Cohen and Stephen A. Merrill, eds. Washington, D.C.: National Research Council, pp. 285–340.

Willig, Robert D. and John P. Bigelow. 2004. "Antitrust Policy Toward Agreements That Settle Patent Litigation." *Antitrust Bulletin*. 49:3, forthcoming.