7 Litigation and Enforcement

7.1. Evidence on Litigation
7.2. Remedies for Infringement and How They Matter
7.3. Technical Protection Measures
7.4. Limited Sharing of Copyrighted Works

A message of chapter 2 was that intellectual property does its work by creating exclusion rights to knowledge or information. The problem of enforcement lies at the very heart of this attempt and is anything but trivial. If the innovator chooses intellectual property over trade secrecy, the technical details are generally disclosed. This is a legal requirement for patents. For copyrighted works, the intellectual property is generally visible on the surface, hence vulnerable to being copied. This is especially true of movies, music, software and books distributed in digital form, although some of those are amenable to technical protections, as discussed in this chapter.

If a protected work is used without authorization of the rightholder, then the rightholder is entitled to remedies. The first step is to prove that the rights have, in fact, been infringed. In legal jargon, the plaintiff must prove “liability”, a nontrivial step. For patents, where the protected subject matter is divided into claims, infringement entails that every element of at least one claim is embodied in the offending product. Although this sounds stringent, we must bear in mind that the claims were chosen by the patent applicant, and approved by the patent examiner, as the minimum combination of elements that constitute an invention. Nevertheless, there is plenty of scope for legal argument regarding what it takes to infringe a claim.

A dispute over intellectual property usually begins as a claim of infringement. In the case of patents, the alleged infringer often counterclaims for patent invalidity – the best defense is a good offense. The counterclaim can be, for example, that the invention was obvious to someone skilled in the art, and therefore did not embody a sufficient inventive step to warrant a patent.

In 1994, the University of California applied for a patent on a technology developed by its researchers for making web browsing capabilities interactive with websites. The technology allows small interactive programs, so-called plug-ins, applets, and scriptlets, to be embedded into web documents. In 1998, UC was awarded U.S. Patent 5,838,906. In the meantime, it licensed this technology to a start-up firm, Eolas, founded by its inventor. While the patent application was pending, Microsoft Corporation, having previously turned down the opportunity to license the technology, embedded such a technology into Internet Explorer, the web browser that is integrated into the Windows operating system. UC sued for infringement. Microsoft defended itself by claiming that the technology had previously been invented by someone else, and therefore the UC patent should be invalidated. The court rejected this defense, affirmed the validity of the patent,
and awarded damages of $520.6 million. Microsoft promised to appeal. Meanwhile, Tim Berners-Lee, director of the World Wide Web Consortium, weighed in with the opinion that the technology was anticipated in prior art, and, bypassing the court system, convinced the PTO to reexamine the patent. This led to a (contestable) decision of the USPOTo to nullify the patent (Stevenson 2004).

Most intellectual property disputes do not involve sums of that magnitude and do not result in reexaminations of patents. However, a counterclaim of invalidity following a complaint of infringement is typical.

For copyrights, the defense is often fair use instead of invalidity. There is a presumption of validity unless the author copied the work from someone else. In theory, infringement of a copyrighted work should be a straightforward matter of checking the offending document against the original, to see if they are “substantially similar.” However, the defendant may argue that she had never even seen the original, and although her work looks similar, it was independently achieved, hence not an infringement. Depending on the circumstances, this defense can obviously strain credulity. If the defendant claims a fair-use exception, as discussed in chapters 3 and 4, the court must adjudicate whether, as a matter of law, the alleged offense is a fair use.

In the 1990s, when the hottest sitcom on television was Jerry Seinfeld and his three disaffected friends, the Carol Publishing Group published a trivia book, the *Seinfeld Aptitude Test* (SAT), in which readers were invited to test their knowledge of “the trifling, picayune and petty annoyances encountered by the show’s characters on a daily basis.” However, the publisher did not have a license to use the Seinfeld characters or to quote from the episodes. The Seinfeld publisher sued. Carol Publishing argued that they only copied minimal parts of the episodes, and even if not minimal, using the material was a fair use. The court rejected both arguments. It held that the SAT passes the test of being substantially similar to the original work, even if there was no literal copying, and also held that the work was not “transformative” in a way that would entitle it to a defense of fair use. The court awarded damages of $403,000.

Whether or not an alleged infringer prevails in court, a good faith belief in fair use or patent invalidity can avoid the conclusion that infringement was willful. Willfulness is relevant to the penalties that can be invoked. The Copyright Act goes so far as to impose criminal penalties (jail time) if copying is willful and for commercial purposes, although criminal penalties are seldom

---

1 As quoted from the author, Beth Golub, by the court of appeals. An example of the SAT cited by the court of appeals is: Who said, “I don’t go for those nonrefundable deals . . . I can’t commit to a woman . . . I’m not committing to an airline.” Multiple-choice answers: (a) Jerry, (b) George, (c) Kramer.
invoked. Patent law provides for treble damages in cases of willful infringement.

Plaintiffs can also be opportunistic and can try the patience of courts, just as willful infringers can. As we will see in the next section, a nontrivial percentage of patents are litigated, and most litigation involves high-value intellectual property, which can be a tempting target.

After J. K. Rowling achieved success in the late 1990s with her Harry Potter books for children, a writer named N. K. Stouer in Pennsylvania filed an infringement suit, claiming she had previously circulated works with a certain Larry Potter, also skinny, bespectacled, and preadolescent (Kirkpatrick 2001). Like Harry Potter, Larry Potter consorted with characters called “muggles.” The characters are similar enough to suggest copying, but also different. Harry has magical powers, but Larry does not. Rowling’s muggles are ordinary folk, whereas Stouer’s muggles are short, hairless, quasi-human mutants.

Part of Rowling’s defense was that she had never laid eyes on Stouer’s work, and therefore could not have copied it. However that would seem extraordinary, given the similarities. The simple explanation believed by the court was that Stouer had made fraudulent claims, and cooked the evidence. The case was dismissed on that basis, with Stouer forced to pay some of Rowling’s legal fees (Italie 2002).

Like other legal disputes, intellectual property disputes often settle out of court. In a settlement, the litigants agree to withdraw their complaints in return for certain payments and/or licensing provisions. Often a settlement will involve the type of license or joint venture that would occur if the case was litigated and a court found infringement and validity. For example, the alleged infringer may abandon the counterclaim of invalidity, and assign his own rights to the patent holder, in return for money. Or the patent holder may abandon her claim of infringement, and grant a license on reasonable terms. In some cases, the litigants merge, or one firm buys the rights of the other.

One thing the litigants are likely to agree on is that the settlement should maximize their joint profit – for example, by imposing high royalties. The settlement gives the two rivals an opportunity to write a contract that joins their interests. As we saw in chapter 6, licensing with royalties can keep prices high. The profit can then be divided by other terms of license such as fixed fees. However this raises an issue for competition policy. If the court would have found invalidity or infringement, and the disputants would then have been competitors, the settlement is anticompetitive. For this reason, settlement of a lawsuit can come under antitrust scrutiny, just

---

2 For more on patent settlements, see Shapiro 2003; Hovenkamp, Janis, and Lemley 2003.
as ordinary licensing contracts can. Sometimes the litigants ask the court to issue an order called a *consent decree*, containing the terms of settlement. This makes the settlement easier to enforce, which may be beneficial to the parties, but also gives the court an opportunity to exercise oversight as to the settlement’s appropriateness.\(^3\)

There is a strong incentive to settle even if the parties disagree on their prospects and the dispute is legitimate. Litigation is very costly. A patent litigation can cost each litigant between $1 million and $3 million per suit (Graham et al. 2003; American Intellectual Property Law Association, 1999) or about $500,000 per patent claim (Barton 2000). As we will see in chapter 9, the value of most patents is considerably less than these litigation costs. Although national litigations in Europe are less costly, between €50,000 and €500,000 (Graham et al. 2002), they might have to be repeated in different countries. Almost 2 percent of patents are litigated. It is difficult to calculate the same statistic for copyrights, because there is no requirement that copyrights be registered, so it is difficult to identify the population that generates litigation.

The next section summarizes some evidence on litigation. Following that is a discussion of remedies for infringement, and how they affect equilibrium profits, a discussion of technical protection measures as a solution to the enforcement problem, and an evaluation of the threat that arises from unauthorized sharing of copyrighted works.

### 7.1 Evidence on Litigation

One of the things that economists have difficulty explaining is why disputes ever reach the courts. Firms can avoid litigation costs by settling. One of the most expensive parts of litigation is “discovery,” the process by which the litigants are allowed to find things out about each other’s cases – for example, by deposing the involved parties or exchanging the views of their experts who would testify at trial on technical matters. Through the discovery process, the firms refine their beliefs about the likely outcome of litigation. Indeed, settlements often occur during this process. However, it is not because the litigants are certain of the outcome that they have an incentive to settle, but only because they are in close agreement on the prospects. If they agree that the outcomes are equally likely, for example, then they should settle on terms that “split the difference” and save the remaining litigation costs.

There is also a question as to why infringement would ever occur in equilibrium. If in-

\(^3\) The court does not always issue the consent decree proposed by the parties, even when the plaintiff is the U.S. government. For example, the court rejected the first consent decree proposed by the parties in the 1994 Department of Justice antitrust case against Microsoft on grounds that it was not in the public interest. See the appendix to Hoerner 1995.
fringement is tempting due to weak remedies or because an infringer is unlikely to be held liable, that should be known to both parties. They can save money and trouble by licensing to avoid the infringement. If infringement nevertheless occurs, the parties can still do better by settling instead of litigating. In fact, settlement may be even more attractive for infringement disputes than for other kinds of disputes, since it may allow the firms to license in a way that increases their joint profit, as compared to the likely outcome of litigation.

Nevertheless, the sparse evidence on litigation suggests that it is prevalent, at least for some types of patents. The author knows of no systematic evidence about the prevalence and cost of litigating copyrights.

The prospect of wasting resources in litigation makes intellectual property a less effective incentive mechanism, just as losing profit to infringers does. Further, if the profit dissipation falls disproportionately on large and small firms, market structure becomes an important determinant of the incentives to innovate. We would therefore like to know how much profit is dissipated through litigation, and which innovators bear the cost. Of course we would also like to know how much profit is dissipated through unremedied infringement, but that is more difficult. We can at most observe litigation, not infringement.

The most comprehensive evidence on patent litigation is due to Lanjouw and Schankerman (2001, forthcoming) who make inferences germane to the questions just raised from a data set they constructed on litigated and unlitigated patents. This data set includes information about the characteristics of the patent owners, the firms’ patent portfolios, and characteristics of individual patents. Much of the data comes originally from the database of issued patents made available by the USPTO. The PTO assigns Patent Classification numbers to every patent, which designate the technological areas. The number of claims, characteristics of the patent owner such as foreign or domestic, the year of application, and backward citations (the prior patents that are cited) can be read directly from the patent document. The PTO’s data also include identifier numbers assigned to patent applicants, so that the database can be used to link all the patents owned by a particular company – that is, the company’s portfolio of patents. The authors then processed the PTO data into a data set that includes forward citations (the patents that come afterward that cite a given patent), self-citations, and portfolio measures (how many patents are held by a given company in a given technology class).

Litigation data are also included in the PTO’s database. When a patent suit is filed in a district court, the clerk of the court is obligated to report it to the PTO. Although such reporting is unreliable, by comparing the PTO data to data collected by the Federal Judicial Center, Lanjouw
and Schankerman reject the hypothesis that cases are underreported in any systematic way. They use that comparison to adjust for the underreporting. The Federal Judicial Center also gives information on the disposition of cases, such as whether they settle and when. A third data source used by Lanjouw and Schankerman is Standard and Poor’s database on companies that are publicly traded, which allows them to know characteristics of the company such as size. By using this data, they can tell whether litigation risks are higher or lower for patents owned by individuals rather than, for example, publicly traded firms.

It turns out that patent litigation varies widely by industry and size of firm, but is frequent and costly for high-value patents, especially in emerging technologies. The overall litigation rate is about 2 cases per 100 patents, concentrated on high-value patents. This litigation rate seems high enough to dissipate a substantial fraction of the reward for innovating. An earlier study by Lerner (1994) was even more pessimistic. Lerner estimated that 6 in 100 biotechnology patents were litigated.

However, even though the number of patent litigations rose in the 1978 – 1999 period of Lanjouw’s and Schankerman’s study, the underlying propensity to litigate did not increase. The increased litigation is attributable to the changing composition of patents, and to the overall increase in patenting. There was a 71 percent increase in patent grants from 1978 to 1995. Most of the increase in patent suits has been in drugs, biotechnology, and computers and other electronics, which have always been highly litigated and have been increasing as a percentage of total patent grants.

The work also paints a rather bleak picture for small innovators. All else equal, a patent held by a large firm is less likely to be litigated than a patent held by a small firm. Previous studies have turned up other disadvantages to small firms. Lerner (1995) concluded that small firms avoid technology areas where litigation is prevalent, and Lanjouw and Lerner (2001) showed that, in the litigation process itself, preliminary injunctions are used strategically by large firms against small firms.

The result that goes most directly to the question of market structure is that patents held by firms in concentrated markets (where patenting is dominated by only a few companies) are less likely to be litigated. A large patent portfolio reduces the probability that any constituent patent is litigated, other things equal. This effect is even more important for small firms, but one suspects that small firms are less likely to have large patent portfolios.

Other specific findings of Lanjouw and Schankerman include the following:

- The overall rate of litigation is about 19 filed suits per 1000 filed patents. The lowest rates are
in chemicals (about 12), electronics (about 15), and mechanical (about 17). Pharmaceuticals are only a little higher than average, but computers and biotechnology are much higher.

- Corporate owners of patents are less likely to be involved in patent suits than individual owners.

- Litigated patents have more claims than average, more forward citations, and fewer backward citations. An interpretation is that the more valuable patents in the newer technologies are most likely to be litigated.

- Ninety-five percent of filed patent suits are settled before trial.

- Firms with the largest portfolios of patents also have the most highly cited (interpreted as most valuable) patents, but nevertheless, the lowest litigation rates.

- The heterogeneity in litigation burden shows up mainly in the propensity to file suits, but not in the subsequent rates of settlement, or disposition at trial.

- Broader patents (those with more PTO technology classifications) are litigated less than narrower patents.

Given the cost and prevalence of litigation, we can conclude that it constitutes an important modification to the profitability of intellectual property rights, and one that differs across different types of firms and technologies.

As shown by the examples at the beginning of the chapter, litigation also gives courts an opportunity to modify the design of intellectual property law itself. This has been of particular interest after 1982, when the Court of Appeals for the Federal Circuit was instituted to hear appeals on patent-infringement cases. The court was established partly because of dissatisfaction among patent litigants over what they perceived as a lack of technical expertise and consistency on the part of federal judges. Patent disputes are still litigated in federal district courts, but most appeals go to the Federal Circuit.

There is at least some evidence that the Federal Circuit has been more “pro-patent” than appellate courts previously handling patent matters. This is a difficult thing to measure, due to the selection of what cases are litigated, and then what cases are appealed. One should be skeptical of attempts to characterize judges’ proclivities by counting outcomes of litigated cases. Nevertheless, scholars have done so, and the data are of some interest.
Harmon ([1991] 1998) summarizes data about three types of patent appeals. In one group, defendants in infringement suits are still pressing their claims of patent invalidity, trying to reverse the decisions of district courts against them. The appellate court thus has opportunity to invalidate a patent. Not doing so might be interpreted as “pro-patent.” In the second group, defendants are trying to reverse a district court’s finding of infringement. The appellate court has an opportunity to exonerate them. Not doing so might again be interpreted as “pro-patent.” In the third group, plaintiffs who lost their suits in district court are asking the appellate court to reverse a finding of invalidity or to overturn a finding of non-infringement. Doing so might be interpreted as “pro-patent.”

In the first two categories, Harmon (1998) reports that the Federal Circuit has affirmed about 85 percent of district courts’ findings for plaintiffs that patents were valid and infringed, and in the third category, that the Federal Circuit has reversed about 25 percent of decisions against the patent holder. Thus, the patent holder has a higher chance of being successful on appeal at the Federal Circuit than the accused infringer does. The accused infringer only has a one-in-seven chance of being successful, while the patent holder has a one-in-four chance of being successful.

In the 1991 version of his work, Harmon reported the same data for a sample from appellate courts prior to Federal Circuit and found they were less favorable to the patent holder. However, in comparing Federal Circuit decisions to a sample of decisions of prior appellate courts, Lunney (forthcoming) finds the opposite. He aggregates the three types of appeals and considers whether the patent holder “succeeded,” which means that the patent holder obtained, at least in part, the remedy sought, such as injunction or damages. He finds that patent holders succeeded in 33.5 percent of appellate cases prior to establishment of the Federal Circuit, but in only 28.3 percent of cases afterward. There was a spike of patentee successes just after the court was established, which may account for the widespread perception that the court is “pro-patent,” but the success rate then tapered off.

What is perhaps of even more interest is that patentees succeed (or avoid failing) in the Federal Circuit for different reasons than they succeeded previously in other appellate courts. Prior to establishment of the Federal Circuit, three-fourths of patentee failures were because some claim of the patent was invalidated. That rate has dropped to about a third during the Federal Circuit period; the Federal Circuit is invalidating fewer patents. Prior to establishment of the Federal Circuit, only about a quarter of patentee failures were because the allegedly infringing device or process was found to be outside the scope of the patent. As discussed in chapter 5, well-designed intellectual property rights can spur firms in a competitive environment to make rapid progress,
taking the form of sequential monopoly. In the era of the Federal Circuit, this fraction has risen to about two-thirds; the Federal Circuit is narrowing the breadth of patents. Lunney (forthcoming) concludes that these data “leave the indelible impression that the Federal Circuit is deliberately and systematically changing the nature of patent law. . . . By eviscerating the nonobviousness requirement, the Federal Circuit has substantially reduced the level of creativity required to establish a valid patent,” and “the Federal Circuit has also limited patents to a correspondingly narrow scope.”

7.2 Remedies for Infringement and How They Matter

There are two kinds of remedies that courts can mete out, injunctive relief and damages. An injunction is a court order to the infringer to stop using or selling the infringing product. Once the court order is in place, the resumption of sales becomes a criminal act, with the possibility of jail time, rather than the mere payment of money to the rightholder. Intellectual property disputes seldom come to this. However, as already mentioned, the battle over liability and damages can be fierce and expensive. It typically involves the testimony of both percipient witnesses and expert witnesses on both the technical merits and the damages.

Damages serve two purposes: they compensate an infringed rightholder for the loss suffered, thus making sure that the rightholder’s ex ante incentive to invest in R&D is intact despite the possibility of infringement, and they can deter the infringement in the first place. We will say that infringement is deterred if the profit available from infringement is less than the prospective damages that the infringer must pay. The two purposes of compensation and deterrence are alternatives: there is no need for compensation if the prospect of damages serves its purpose of deterring infringement.

There have been two theories of damages in the common law: the infringer must pay the rightholder’s lost profit, or the infringer must disgorge his unjust enrichment. The measure now used in the United States is lost profit, which is aimed at replacing the infringed rightholder’s losses. Unjust enrichment would be aimed at making sure infringers do not profit from their misdeeds. Oddly enough, neither doctrine seems explicitly aimed at deterring infringement, although either may do so, as we will see. The court also has the option to impose additional punitive damages if

---

4 The discussion here is based mostly on Schankerman and Scotchmer 2001, especially in its emphasis on equilibrium profits, as determined by the threat of infringement and damages. See Calabresi and Melamed 1972 for the key treatment distinguishing property and liability rules, and Conley 1987 for a good description of how the judicial treatment of remedies has evolved. See Heath, Henkel, and Reitzig 2002 for a discussion of how these rules have been developed in different countries.
the patent infringement is deemed willful, and criminal penalties are available under copyright law.

7.2.1 Remedies: Product Patents

Figure 7.1 shows the difference between lost profit and unjust enrichment when a product patent is infringed. The left diagram shows the rightholder’s per-period monopoly profit, absent the infringement. The right diagram shows the rightholder’s oligopoly profit when competing with the infringer, and when the two firms share the market as shown. Let $\pi_I^P$ and $\pi_I^C$ represent their respective profits if they compete, where the subscript on $\pi$ denotes (I)nfringement and the superscripts denote (P)atentholder and (C)ompetitor (infringer). Competition lowers the market price, and, in addition, the rightholder loses market share. The patent holder’s lost profit would be $M^M - \pi_I^P$, and the infringer’s unjust enrichment would be $\pi_I^C$. Since the monopoly profit $M^M$ is always larger than the sum of the two firms’ oligopoly profits $\pi_I^P + \pi_I^C$ (else the monopolist would have chosen a lower price), it follows that $M^M - \pi_I^P > \pi_I^C$. Thus, due to price erosion, the lost-profit rule yields larger damages for the rightholder than the unjust-enrichment rule does.\textsuperscript{5}

Suppose now that the purpose of the damages rule is to deter infringement. The lost-profit rule will do so, since $M^M - \pi_I^P > \pi_I^C$. But the unjust-enrichment rule might also deter infringement. Since the infringer is made to disgorge the entire ill-gotten gain, he is indifferent between infringing and not, and it is reasonable to assume that he will be deterred, especially if there is any chance that punitive damages will be imposed.

In the case depicted here, the lost-profit rule is better for the infringed rightholder when

\textsuperscript{5} Commentators sometimes claim that the unjust-enrichment measure is larger than the lost-profit measure whenever the infringer is more efficient at working the patent than the patent holder. That view does not take account of licensing. If the infringer is more efficient, the likely (efficient) equilibrium outcome would be licensing by the patent holder, not production by the patent holder. See the next subsection, which discusses how to measure lost profit when licensing would alternatively be the equilibrium outcome.
deterrence fails, since he is “made whole.” This is not true under the unjust-enrichment rule. However, if infringement is deterred in equilibrium, this argument may not matter. It does not matter what would happen in the out-of-equilibrium event of infringement, since an out-of-equilibrium event is not expected to occur. Nevertheless, if both rules deter infringement, there is no harm in choosing the one that makes the rightholder whole in the event of infringement.

Even so, there is a problem knowing how to calculate lost profit. In the circumstance depicted by figure 7.1, experts would have to estimate the demand curve and then use it to predict the price that would have prevailed, absent the infringement. The court would have to make a judgment as to the equilibrium in the left-hand diagram, and calculate the lost profit $\pi^M - \pi^P$. Naturally courts are reluctant to do this. It is tempting to use a more limited notion of lost profit, which assumes that the prevailing price with infringement, $p^O$ in figure 7.1, would prevail even absent the infringement. Lost profit is then calculated as lost sales at that price. In figure 7.1, this calculation would lead to a damage estimate of $\pi^O$. This method avoids the problem of predicting the price that would have prevailed, absent the infringement, but can significantly underestimate the loss. Another alternative is simply to take account of price erosion but not the loss in sales. This method also underestimates the true loss. The estimate would be $\pi^P \times (p^M - p^O)/p^O$. The underestimation matters for both compensation and deterrence. The infringed is not compensated, and the infringer may not be deterred.

Licensing changes this story in its details but does not change the conclusions, provided the true lost profit can be imposed as damages. Suppose that the rightholder is infringed but that marginal cost is increasing instead of constant, as in the model of section 6.1. There are four possible counterfactuals to infringement in that situation. One possibility is that the rightholder would have supplied the whole market inefficiently, using a single production facility. Another possibility is that the rightholder would have built a second plant and supplied the whole market efficiently. A third possibility is that the rightholder would not have produced at all, licensing two firms instead. A fourth possibility is that, absent the infringement, the rightholder would have supplied part of the market and would also have licensed the infringer.

With all these counterfactuals, which should the court use? And how should it be applied? With constant marginal cost, it was natural to assume that the rightholder would have supplied the market in a single production facility. The only problem was to estimate demand and find the profit-maximizing price. Here, the task is first to figure out the right counterfactual, and then to calculate the lost profit. Some of the counterfactuals involve licensing and some do not. Some involve efficient production and some do not. The counterfactual lost profit may include both lost
sales revenue and lost licensing revenue.

The view taken by Schankerman and Scotchmer (2001), followed here, is that lost profit should be calculated by reference to an equilibrium that “should have occurred.” This perspective goes a long distance toward cutting through the complexities of the competing counterfactuals. At a minimum, an equilibrium should be efficient. We can eliminate the counterfactual with inefficient production in a single facility, because that will not be profit maximizing for the rightholder. The rightholder will either build a second plant or license. Among the other three options, we saw in section 6.1 that if the licensor can use royalties and fixed fees, she can make the same profit by licensing two firms as by supplying the market herself, producing in two production plants. With enough permissible terms of license, she can also do as well by licensing a single firm and supplying part of the market herself. We can therefore take this maximum profit as the reference point, without specifying which of those three efficient options she would use. Of course, in one of those outcomes the profit would be earned entirely through licensing fees, in another it would be earned entirely through sales revenues, net of costs, and in the third, it would be earned through both sales and license fees. But all would give the same profit to the rightholder.

Suppose, then, that absent the infringement, the patent holder would have earned profit \( J \). In the notation of section 6.1, \( \pi^J = q^*p(q^*) - 2 \int_0^{q^*/2} \gamma(\bar{q})d\bar{q} \), and \( \pi^J \) is larger than the profit that the rightholder would have earned as a monopolist with a single production facility. As before, we can let \( \pi^P_I \) and \( \pi^C_I \) represent the patent holder’s and competitor’s (infringer’s) profits with infringement. Since \( \pi^J > \pi^P_I + \pi^C_I \), it remains true that the patent holder’s lost profit \( \pi^J - \pi^P_I \) is larger than the infringer’s gain \( \pi^C_I \). Thus, the lost-profit rule will deter infringement.

For this argument, we have not had to disaggregate the lost profit into lost sales net of costs and lost licensing revenue. However, that is where it gets difficult from a practical point of view. Although it is firmly established in legal doctrine that lost profit can be lost royalty (“reasonable royalty”), there are no clear guidelines for calculating it. Since the putative equilibrium never happened, the court cannot refer to the licensing arrangements or sales revenue that underlie \( \pi^J \), and in any case, there are various licensing arrangements that could do so, including no licensing at all. To reduce the degree of speculativeness, courts may feel constrained to choose between a measure of damages based solely on lost sales and a measure of damages based solely on lost royalties.

Thus, although the lost-profit doctrine grants enough penalties to deter infringement even without punitive damages, it is difficult to calculate the benchmark level of profit. Fortunately, infringement may be deterred even if lost profits are underestimated. That is because, in the case
of product patents, infringement dissipates profit in total, as well as the rightholder’s profit. The profit dissipation is itself punitive, so that the minimum damages required to deter infringement are less than the lost profits. If lost profits are measured accurately, infringement is “overdeterred.”

That simple story changes when licensors and their licensees are not competitors in the market, such as in the licensing of research tools.

7.2.2 Remedies: Licensed Research Tools

As we stressed earlier, the difficulty in calculating lost profit is that it can only be calculated with reference to some hypothetical equilibrium that should have occurred but did not occur. Since profit is an equilibrium concept, “lost profit” is also an equilibrium concept. But the problem of calculating the equilibrium that should have occurred is compounded when licensing is at stake, because the license fees at which the infringer “should have licensed” are also part of the equilibrium that did not occur. This problem was avoided in the case of patented products by referring to a benchmark level of profit that could be achieved in various ways, with or without licensing. That dodge will not be possible in what follows.

The focus here is mainly on research tools. Tools have become particularly important in biotechnology, where the process of developing new drugs or other products requires (1) a suitable host, like bacteria, tobacco plants, or (for agricultural products) a germplasm, (2) techniques for inserting foreign genes into the host, (3) genetic sequences that code for the desired trait or protein, (4) genetic sequences that code instructions to the host, such as suppression of other genes, or expression, and (5) techniques to extract the product without killing the host. The developer of a bioengineered product will not typically own the intellectual property that gives access to all five pieces, and must therefore license.

In this circumstance, the lost-profit doctrine leads to an indeterminacy in licensing fees that may undermine the incentive to invest in research tools. This is due to a circularity in the definition of lost royalty.

Suppose that the license fee is some number of dollars, $L$, and the damages for infringement under the lost-profit doctrine would be some number $d$. Looking at the court’s calculation of damages under the lost-profit doctrine, damages should satisfy $d = L$. The maximum that a potential licensee would pay for a license is the damage he would otherwise have to pay, $d$. Further, the minimum that a licensor would charge is also $d$, since she could get $d$ in a court award afterward. Thus, it must hold that $L = d$. But what determines the mutual values of $d$ and $L$? There is nothing to nail it down, except that the mutual value must be less than the licensee’s willingness to pay. By
Willingness to pay for license, \( w(\theta) \)
Licensee, \( \theta \)

Figure 2: A market for licenses

Definition, a licensee would not pay more than his willingness to pay, and such a fee could therefore not be lost profit.

In short, the whole line of reasoning is circular, and hence not very satisfying. It means that the profitability of the licensed innovation is indeterminate, and the possibility of low rewards may discourage innovation.

To elaborate this argument, suppose that the market for the research tool is anonymous, in the sense that the licensor cannot price discriminate according to the user’s willingness to pay. The market demand for such licenses is shown in figure 7.2. Figure 7.2 might depict, for example, the market for PCR licenses. PCR is a chemical procedure, extremely useful in biomedical research, that allows DNA to be replicated easily for experimental use. For such tools, each user’s willingness to pay is either determined by the convenience of not using the next best alternative, or by the value of the research program it enables.

At the original demand curve, or willingness-to-pay curve (the solid line), the licensor’s best option is to charge the monopoly price, \( L^M \). That price should be self-sustaining. An infringer would have to pay damages \( d = L^M \), since that is the prevailing price, and is thus the lost profit when an infringement occurs.

Suppose, though, that demand shifts to the dotted line in figure 7.2. The monopoly price that the licensor would then like to charge is \( \hat{L}^M \). Can the price be raised?

Due to the self-reinforcing nature of damages and license fees, increasing the price will be difficult once it is established, with the consequence that the prevailing price can be somewhat arbitrary. If the licensor suddenly tries to charge the price \( \hat{L}^M \) to a new licensee, the licensee may well reason that infringement is a better option. The court will charge damages of \( d = L^M \), which is the observable prevailing price and hence the court’s assessment of lost profit. Unless the licensor
can find a way to change the price for all licensees simultaneously, it will be difficult to adjust the price.

The same argument can be made without reference to a shift in demand. If, for any unspecified reason, the licensor finds herself charging a price that is lower than the most profitable one, she will have trouble adjusting it, due to the self-reinforcing nature of the damage measure and license fees. Thus, there is nothing to pin down the equilibrium license fee and damages, and they may be somewhat arbitrary. A licensor can get locked into an equilibrium with price lower than the most profitable price.

7.3 Enforcement of Copyrights by Technical Protection Measures

Intellectual property rights are inherently difficult to enforce, since they are exclusive rights to “information,” and information has the character of a public good. It can be copied or used freely once it exists. For manufactured objects, the threat of infringement is kept under control because infringement does not take place at the level of individuals, but rather at the level of firms. It requires a production and distribution facility, which are public enough to be detected.

Intellectual property law provides a legal means by which inventors can exclude unauthorized users from using their creations. There are also technical means, such as encryption and copy controls. The technical means may obviate the need for legal means, but they are not equivalent.

Technical protection measures will be particularly important if the legal means of exclusion fail — that is, if copyrights or other intellectual property rights are not enforceable. Technical protections have been more widely deployed in the copyright realm than the patent realm, but even in the patent realm, there is at least one important example, the so-called terminator gene. This is a gene developed by Monsanto that, if inserted into genetically modified crop seeds, will prevent a second generation of seeds from germinating. As a consequence, the farmer cannot save seed for planting the next crop. Saving seed would generally be prohibited by license for patented genetically modified crops in any case, but such a prohibition is hard to enforce. The terminator gene solves the enforcement problem but is controversial for other reasons, such as its potential for finding its way into other seeds. Wide introduction of the terminator gene could have disastrous consequences.

Technical protections have been more widely introduced for so-called digital content such as movies, music, computer software, and computer games (see Samuelson and Scotchmer 2002). If these products are distributed on CDs, they are easy to copy without detection. Of course, print books and music distributed on audiotapes have long been vulnerable to copying. Such copying
has not undermined their market in the past, because the inconvenience of copying has been more costly for most consumers than the price of a legitimate copy.

The threat to copyright holders has evolved as technology has evolved. In the 1970s, when videocassette recorders were introduced, movie producers and other vendors of broadcast and cable content feared that home videotaping would undermine their ability to collect box-office fares or rent videos. Did the viewer have a right to tape movies for personal use? This issue went back and forth in the courts in the 1970s and 1980s.

To some extent analog tapes have self-limiting copy potential, in that the quality of each successive copy degrades. Thus it is difficult to make unlimited copies of copies, which to some degree limits the threat of copying. In contrast, digital content can be copied faithfully. When digital audiotapes (DATs) were introduced in the mid-1980s, content providers feared that the ability to make faithful copies of copies would lead to widespread infringement. By 1992 they had persuaded Congress to adopt a copy management solution in the Audio Home Recording Act. This act mandated that DATs and DAT players must contain a Serial Copy Management System, a technological modification that allows the making of copies, but only with some degradation, similarly to analog tapes.

The 1980s saw technical protections fail in the market. In the late 1980s, software for personal computers was sold with a one-installation feature. Users (hence vendors) found this annoying, because any glitch in installation would render the software unusable. Further, users could not make back-up copies of the software, or install it on both their home and office computers, as they might think “fair.”

Ultimately the one-installation feature disappeared from the market. In any case, today’s computers could copy one-installation disks before installation, thus undermining the system. The modern incarnation of the one-installation feature is to download software directly from the Internet to the user’s computer. Downloading is a double-edged sword. For computer software, it is a convenient and controlled way to sell a single installation without circulating a disk. But for content such as music, it can facilitate the distribution of illicit copies, as in the Napster case discussed later.

Other solutions to the copying of digital content are encryption and watermarks. These are attempts to make the digital content uninterpretable or inaccessible without use of a code key. Access to the key must be authorized by license or sale, often with the sale of the content itself.

---

Of course, one might ask what is different about selling access to a key rather than access to the content. If the content can be copied, so can the key. Further, the key is typically a password, and if the user really wants to overcome it, he can use another computer to try passwords seriatum. A more reliable version of the key system is call-home authorization, a version of the direct download. Instead of downloading and installing the software from the CD alone, the user logs on to a web site in order to pay the vendor and receive an authorization code for the software received on CD.

A watermark, by analogy with watermarks on paper stationery, is a piece of software code embedded in a program. It may have two purposes: First, if illicit copies of the software circulate, the watermark can identify the original buyer or licensee of the copy that is circulating. This may or may not be useful, depending on whether the original buyer or licensee can be held liable. The second purpose works in conjunction with code embedded in an authorized platform required to use the digital content. For example, the movie industry has developed digital versatile disks (DVDs), which are protected by a technology called the Content Scrambling System (CSS). CSS authorizes access by matching a code embedded in disks to a code embedded in DVD players. Manufacturers of DVD players must license the CSS system, which must itself be protected for the system to work. The sound-recording industry tried to implement a similar system, called Secure Digital Music Initiative (SDMI), by which digital sound recordings would embody special authorization signals to be detected by authorized players. For various reasons, that system has not succeeded.

These systems shift the enforcement problem from protecting the content to protecting the players. In their most extreme implementation, the content would not be protected at all. The entire enforcement burden would be moved to players. Protecting players is not a perfect substitute for protecting content, however, since it is difficult to distribute the revenue from players in a way that rewards content according to its popularity. Another problem with this solution (a problem for antitrust authorities, but not necessarily for content providers) is that the joint ownership of the players’ authorization technology might allow an implicit collusion among content providers that would be avoided by protecting content directly (Scotchmer and Park 2004). And at the technical level, the system might not work because the technical protection system might be reverse engineered, thus opening the player market to unlicensed entry.

The message in the preceding account is that technical protections are so far limited in their ability to protect content. There is probably no perfect way to exclude unauthorized users from access. At best, technical protection measures can make it costly to get at the content, for example, by having to circumvent a password, encryption system, or watermark. But even encrypted digital content is vulnerable to copying. In order to use the content, it must at some
point live unencrypted in a computer’s memory. At that point, even if no other, it is vulnerable. In short, according to what the market has produced so far, getting around technical protection systems is simply a matter of being willing to bear the cost.

In response to the perceived inadequacies of technical protections, Congress enacted the Digital Millennium Copyright Act (DMCA) in 1998. Among other provisions, and with some exceptions, the DMCA prohibits circumventions of technical protection systems.\footnote{For a discussion of the limitations on this threat, and how they had been applied by 2002, see Samuelson and Scotchmer 2002.} That is, it creates legal protections for technical protections.

A logical question is why legal protections for technical protections are likely to work, given that legal protections for digital content are assumed not to. One obvious answer is that, in addition to creating rules to protect the protections, the DMCA imposes ferocious penalties for violating the rules. Under the DMCA, it is a criminal act to circumvent a technical protection measure for a commercial purpose. That is, the guilty citizen can expect jail time rather than the financial penalties discussed earlier. Criminal penalties can apparently apply even if the technical protection is trivial (such as guessing a two-character password), although no such cases have been brought. Under the No Electronic Theft Act of 1997, it is also possible to incur criminal liability for direct copyright infringement, even without a circumvention, and for copyright infringements with no commercial value. Both of these acts have exceptions that may exempt the inadvertent infringer.

Even if not entirely effective, technical protections can have important economic consequences. These include

1. All technical protection measures involve wasted costs, at least as compared to an idealized world in which users automatically respect intellectual property rights. The waste is compounded if users and providers get caught in a measures-on-countermeasures war with every attempt at technical protection defeated with a countermeasure, leading to escalation on both sides.

2. If technical protection measures become truly effective, they will likely undermine fair uses of digital content, since fair users are excluded from access, just as infringers are. As mentioned in chapter 4, there are solid policy reasons to grant fair use of copyrighted content. These may be sacrificed if technical protection measures become truly effective.

3. If technical protection measures are truly effective, there is no reason to think that the
duration of protection will accord with the provisions of intellectual property law. If an inventor can protect his invention for twenty years using a technical protection measure, then why not thirty? That is, he can protect it as a trade secret much longer than it would be protected with the legal protection of a patent.

4. Technical protections can avoid the disclosure requirement that would be required in the case of patents. To assess the economic consequences, one must have a theory for why disclosure is required; see chapter 5. To the extent that disclosure aids rivals, technical protection aids the rightholder, but may it inhibit innovation in aggregate.

Thus, limitations on intellectual property rights may vanish if the rights are protected technologically rather than legally. This observation should give pause to the enthusiast who thinks that legal protection should give way to technological protection. To the extent that technical protection systems are truly effective, the issues of designing intellectual property discussed elsewhere in this book may become moot.

However, because technical protections are expensive to implement, and because some users have the ability to circumvent them for personal use without detection, it is not obvious that technical protections will lead to higher prices for users. Even though Congress has legislated severe penalties for individual acts of circumvention as well as for commercializing circumvention tools or commercializing pirated works, individual acts of circumvention are hard to detect. If users will circumvent the protection system when the cost of circumvention is lower than the price, the threat of circumvention will have a moderating effect on the pricing strategy of vendors, as follows.

Index the strength of protection by $e$, and for simplicity, assume that $e$ also represents the cost of circumvention. The vendor cannot sell units at a price higher than $e$. By choosing the strength of protection, $e$, the inventor thus chooses the maximum feasible price.

Suppose that the demand at price $p$ is $q(p)$, and that the marginal cost of supplying units of the content is zero. Then the monopoly price with perfect legal enforcement would be the $p^*$ that maximizes $pq(p)$.

The profit available in the market with protection $e$, as depicted in figure 7.3, can be written

$$\Pi(e) = \begin{cases} p^*q(p^*) & \text{if } e \geq p^* \\ eq(e) & \text{if } e < p^* \end{cases}$$

The vendor will not raise price higher than $p^*$, even if the protection system is strong enough to permit it. Doing so will not increase profit. Thus, the profit function $\Pi(e)$ flattens out for $e \geq p^*$, as shown. If the protection system will not sustain a price as high as $p^*$ due to the threat of
Figure 3: Optimal technical protections

circumvention (that is, if $e < p^*$) then the vendor will price as high as possible while avoiding circumvention.

Let $K(e)$ represent the cost of implementing a protection system of strength $e$. The cost of protection may account for the cost of updating an encryption or watermark system to stay ahead of the hackers, or it may simply be an upfront cost. The cost curve $K$, assumed to be increasing, is also drawn in figure 7.3. The optimal level of protection is where the distance $\Pi(e) - K(e)$ is greatest, which is $e^\ast$. The optimal level of protection is less than the level of protection that would support the monopoly price, $e^\ast$. By reducing protection marginally from $e^\ast = p^*$, the vendor reduces the costs of implementing the system, but there is very little impact on profit.

Thus, a reliance on technical protection has a moderating influence on price, which reduces per-period deadweight loss. Weighed against this reduction in per-period deadweight loss, however, is that the protection system is costly, and it can continue indefinitely, rather than terminating at a stipulated time, as intellectual property protection does. Taking account of these offsetting effects, it is not obvious whether the threat of circumvention reduces social welfare.

Surprisingly, it is possible that a reliance on technical protections can benefit both consumers and vendors. This follows from the arguments in chapter 4, which showed that consumers may be better off if the intellectual property right is longer, provided the price in each period is lowered enough so that discounted profit stays the constant. The same reasoning carries over to technical protections. The technical protection regime may prolong protection, but also leads to a lower price.
Internet service providers (ISPs) are also important actors in the enforcement environment, possibly as contributors to infringement, and possibly as allies in the enforcement wars. In copyright law, contributory infringement is a willful attempt to encourage or facilitate infringement by someone else, and the facilitator becomes liable. ISPs bear such a risk if the web sites they host contain pirated content. As a consequence, they persuaded Congress to write a provision in the DMCA that gives them a safe harbor unless informed of the violation.

A general economic principle is that monitoring for enforcement should be done by the actor who can do it at least cost. This may be either the content providers or the ISPs. Content providers have a clearer sense of what constitutes a violation and have an incentive to inform ISPs. However, they will likely be aggressive in claiming infringement, and thus put pressure on ISPs to evict sites that might be legitimate fair users. The ISPs will be under the burden of making legal judgments as to potential liability, whether they themselves are the monitors, or simply assume liability when informed.

7.4 Limited Sharing of Copyrighted Works

The ease and accuracy of copying digital content have made providers fearful of large-scale piracy. These fears have been fueled by episodes of widespread copying in Asia, as well as well-publicized technologies for unauthorized sharing that have invaded the American market, such as the now defunct Napster. Napster was a database that allowed users of music to find other users in possession of music they wanted. After the two users found each other, one user could download music from the other user’s computer. Napster was shut down after a lawsuit by music producers, on grounds of contributory infringement. Similar websites are now operating, and several charge fees that can be passed back to content providers.

In the “old” economy, wide-scale infringement was detectable because it typically required a production facility. A production facility is hard to obscure. In contrast, digital content can be copied by anyone with the capacity to burn a CD or download from the Internet.

But exactly what threat does this pose? Perhaps the most important lesson of Napster is not that copying is easy, but that the music industry succeeded in shutting it down. This is again because the infringement was taking place on a scale that was hard to miss.

Under the first-sale principle of copyright law, buyers can dispose of their lawfully acquired copies in any way they want to, including resale or lending. This right extends to digital content such as CDs, provided the CD is borrowed as a physical object and not copied. The threat perceived by content providers is that the CD is likely to be copied instead of borrowed. Most
of us know people who have engaged in some form of illicit sharing through copying. However, if such sharing is hard to detect, it is presumably because the sharing circle is small. The kind of sharing that takes place at the level of friends is limited in scope.

Bakos, Brynjolfsson and Lichtman (1999) argued that, if sharing is limited in scope, content providers might actually make even more profit if they anticipate sharing than if not. They will price in a way that anticipates the sharing. To see why this might be true, suppose there are six customers, with willingness to pay (1,1,4,5,6,7). If the content is sold at a monopoly price, the most profitable price is 4, yielding profit 16. There are four buyers with willingness to pay greater than 4.

Suppose instead that these buyers coalesce into sharing groups of size two. In that case, the willingness to pay of each group can be treated as a single number, the assumption being that there will be some sort of side payment between the friends that resolves their disagreements about what to buy. The following are two ways that the same aggregate willingness to pay can be distributed in groups of size two:

\[(5, 6, 13) \quad (6, 7, 11)\]

These divisions into sharing groups yield different market demand and call for different profit-maximizing prices. The profit-maximizing prices are, respectively, 5 and 6. In both cases, all three groups will buy the content, yielding profits 15 and 18. The total profit is lower than 16 in the first division of consumers and greater in the second. Thus, sharing can either increase or decrease profit, depending on the groups that form.

However, an important criticism of this line of reasoning is that it takes group formation to be exogenous. Of course it is true that much sharing takes place informally in groups that are thrown together for other reasons, such as living on the same floor of a dormitory. Nevertheless, such groups are not random. In fact, the members are likely to have a lot in common. This is reinforced by unspoken protocols of sharing, such as that each member must occasionally subscribe to a magazine that everyone likes or buy a piece of software that they all want. A member who contributes nothing that the others want would soon be evicted from the group. Thus, although sharing groups probably do not form strategically for the sole purpose of sharing intellectual property, they are also not random.

Suppose, then, that we study the opposite extreme, that groups form for the dedicated purpose of sharing software or content, and that there are informal side payments within the group to equalize the costs. Some of these side payments may take the form that different members buy
different contributions. It is shown by Scotchmer (a) that, in this situation, there is neither a profit advantage nor a profit disadvantage to sharing, provided the maximum size of the sharing group that can escape detection is fixed.

To see this, suppose that each consumer has willingness to pay either \( a \) or \( x \) for each CD, \( a > 2x \). An agent of type-\( v_1 \) has high willingness to pay for jazz and low willingness to pay for classical, and for type-\( v_2 \), it is the other way around. The two types are equally represented in the population. That is, their willingnesses to pay are

<table>
<thead>
<tr>
<th>type of person</th>
<th>( v_1 )</th>
<th>( v_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>WTP classical</td>
<td>( x )</td>
<td>( a )</td>
</tr>
<tr>
<td>WTP jazz</td>
<td>( a )</td>
<td>( x )</td>
</tr>
</tbody>
</table>

Suppose first that the monopolist sellers of the classical CD and jazz CD sell separately to single buyers. Since the problem is symmetric for the two sellers, consider the profit of the classical monopolist only. There is no point charging any price between \( a \) and \( x \). At price \( p = a \), half the consumers buy, so profit per capita is \( a/2 \). If the monopolist sells at price \( p = x \), then everyone buys, so the expected profit per capita is \( x \), but that is less profit than selling to half the buyers at the higher price, since \( a/2 > x \).

Now suppose that the monopolist sells to pairs of buyers who share the CD – everybody has a friend. At given prices, the pair will buy the CD if their combined willingness to pay exceeds the price. If one of them values it more than the other, they will find some way to compensate within the sharing groups, so it evens out.

There are three types of groups: two people with tastes \( v_1 \), two people with tastes \( v_2 \), and mixed groups with one person of each type. In a group with one person of each type, the joint willingness to pay is \( a + x \) for each type of CD. At prices for classical and jazz \((p_c, p_j) = (a + x, a + x)\), they will buy both CDs. If all the consumers are matched in that way, then the total profit per capita is \((a + x)/2\), which is larger than the profit when selling to single buyers at either \( p = a \) or \( p = x \). This is the type of phenomenon pointed to by Bakos, Brynjolfsson and Lichtman (1999), who point out that if the willingnesses to pay for each CD are negatively correlated within the group, then the proprietors can actually make more money, not less, when consumers share. The buyers with low willingness to pay are not left out of the market, because they are paired with high willingness to pay, and the combined willingness to pay may be enough to induce a sale.

Matching the consumers in mixed groups maximizes the profit for the proprietors. However, there is no reason to think that groups will form in this way. Suppose instead that they form randomly, such as according to who happens to be neighbors. Then the distribution of willingnesses to pay of the groups is given by:
The most profitable prices are still \((p_c, p_j) = (a + x, a + x)\).

However, it is rather odd to assume that the pairs will be assembled in any way that is not rational for the buyers, conditional on the prices. Certainly if the prices are \((p_c, p_j) = (a + x, a + x)\), it is best for consumers to assemble according to shared tastes. Suppose that \(v_1\) consumers pair up with other \(v_1\) consumers, and the same for \(v_2\). Then at prices \((p_c, p_j) = (a + x, a + x)\), the pair of type-\(v_1\) consumers will not buy the classical CD but will buy the jazz CD. Instead of getting zero surplus from their purchases, as the mixed groups do, they get a positive surplus: 

\[
2p_j - 2a = 2(a + x) - 2a = 2x > 0.
\]

In fact, whatever prices are offered by the proprietors, the consumers will pair up optimally, which means in this simple example that they always pair up with someone who shares their taste vector, \(v_1\) or \(v_2\). (In a more general model, where all the consumers can have different tastes, the optimal way of forming groups will depend on the prices (Scotchmer, forthcoming). Further, if they always pair up optimally conditional on the prices, the proprietors end up with exactly as much profit as if they sold to individual consumers. The proprietors will sell at higher prices, but only high enough to account for the fact that they are selling to groups of consumers rather than to individuals.

To see this neutrality result in the example, let the prices \((p_c, p_j)\) be arbitrary, and consider four agents, two of each type. The total consumers’ surplus in two mixed groups is smaller than the total consumers’ surplus available if the same consumers split into two homogeneous groups, one of which has two \(v_1\) consumers and the other of which has two \(v_2\) consumers.

Consider the classical CD. (The argument for jazz is the same.) If \(p_c < 2x\), then all the consumers will buy the CD regardless of how they are grouped. If \(2x < p_c < a + x < 2a\), then the mixed groups will buy the CD, with total consumers’ surplus per group of \(2(a + x) - 2p_c\). The \(v_2\) groups will buy the CD but not the \(v_1\) groups. The total consumers’ surplus with homogeneous groups is \(2a - p_c\), which is larger than with mixed groups. If \(a + x < p_c < 2a\), then the mixed groups do not purchase the CD and receive no consumers’ surplus from classical CDs. However, the homogeneous \(v_2\) groups will buy it and receive positive consumers’ surplus \(2a - p_c\). This shows that grouping the consumers according to taste cannot reduce the consumers’ surplus, and may well increase it.
References and Further Reading


Eolas Technologies v. Microsoft Corp. N.D. Ill., No. 99 C 626, 8/11/03.


Journal 111:1575-1663


