The Essentiality Test for Patent Pools

Richard J Gilbert
THE ESSENTIALITY TEST FOR PATENT POOLS

RICHARD GILBERT*

I. Introduction

Antitrust policy for the pooling of patents and other intellectual property rights has undergone a dramatic transformation since the first cases were decided at the beginning of the twentieth century. This transformation generally reflects developments in economics that provide a better understanding of the characteristics of patent pools that warrant antitrust scrutiny. The change, however, was slow, and the US antitrust agencies did not clarify their enforcement principles with respect to patent pools until the publication by the Department of Justice (DOJ) and the Federal

* Emeritus Professor of Economics and Professor of the Graduate School, University of California at Berkeley. I am grateful to Michael Katz and Carl Shapiro for helpful discussions and to Harry First for comments on an earlier draft.

The DOJ/FTC IP Report concludes that a patent pool is unlikely to raise antitrust concerns if:

- the pool is limited to patents that are essential to implement the standard;
- the pool grants non-exclusive licenses that do not prevent licensees from developing alternative technologies;
- patentees grant non-exclusive licenses to the pool and retain the right to license their patents separately outside the pool; and
- licensees are required to grant back nonexclusive licenses to use patents they hold that are essential to comply with the technology.

The definition of an essential patent can take different meanings. A patent may be technically necessary to produce a DVD, while another patent may be necessary only to produce a DVD that displays subtitles. A patent may not be technically essential because it has alternatives, but the alternatives may be so inferior as to render them very distant substitutes.

Patents that are technically essential or that have no close substitutes are comple-ments to implement a standard or to make or use a product that employs the patented technology. Economic theory implies that pooling complementary patents and licensing the patents by a single entity is pro-competitive relative to individual licensing. With individual licensing there is the risk of inefficiently high total royalties if the patentees charge running royalties, because each patentee would ignore the effect of its royalty on the demand for other patents necessary to implement the technology. A single entity that licensed all of the patents would internalize this effect and have an incentive to charge a lower total royalty for the pooled patents than the sum of all the royalties charged by individual licensees.² Furthermore, if individual patentees fail to coordinate their licensing, this could delay the implementation of the technology. A single licensing entity has an incentive to coordinate the pooled licenses and would minimize licensing delays. Pooling the patents also reduces the transaction costs of negotiating licenses.

Much of the effort on the part of patent pools to comply with DOJ and FTC enforcement principles has gone into ascertaining whether the pool is limited to patents that are essential to use the licensed technology or implement a standard.

---


For example, in its 1997 Business Review Letter for the proposed MPEG-2 patent pool, the DOJ took comfort in the fact that ‘The Portfolio combines patents that an independent expert has determined to be essential to compliance with the MPEG-2 standard; there is no technical alternative to any of the Portfolio patents within the standard.’ Furthermore, the DOJ noted approvingly that the pool would be limited to ‘technically essential patents, as opposed to merely advantageous ones,’ and that an independent expert would play a continuing role to ensure that the pool is limited to patents that are technically necessary to implement the standard.

The DOJ considered a more expansive definition of essential patents in its 1998 and 1999 Business Review Letters for two patent pools related to the DVD-ROM and DVD-video formats. The first DVD pool involved three licensors and is sometimes called the 3C DVD pool. The second involved six licensors and is sometimes called the 6C DVD pool. The 3C DVD pool broadened the concept of essentiality to encompass patents that are economically unfeasible as well as patents that are technically essential because they are inevitably infringed by compliance with the specifications. The 6C DVD pool considers a patent to be essential if it is ‘necessarily infringed’ or if there is ‘no realistic alternative’ to it.

The DOJ remarked that these definitions of essential patents are broader than the definition of ‘technically essential’ that applied to the MPEG-2 patent pool and noted that they raise competition issues. In particular, the DOJ expressed concern that the pool could be a price-fixing mechanism if it includes patents that are substitutes for each other. The DOJ also noted that the pool might foreclose competing patents if it includes a patent with a substitute but excludes the substitute. Manufacturers might choose not to license a competing patent that is outside the pool because the pool would offer a package license that includes one of the substitute patents. The DOJ ultimately concluded that the broader definitions


2. Ibid.

3. DVD stands for digital versatile disc.


8. Ibid.


10. Ibid.

11. See ibid.
of essential patents offered reasonable protections against anticompetitive outcomes.\footnote{12}

A 2008 DOJ Business Review Letter related to a standard for ultra high frequency radio frequency identification (UHF RFID) affirmed the DOJ’s perspective on the types of patents in patent pools that do not raise antitrust concerns.\footnote{13} The DOJ noted approvingly that the RFID Consortium, which licenses patents to implement the Generation-2 standard for UHF RFID devices, would limit the pool to essential patents and would exclude from the pool any patents that are found to be invalid. The definition of an essential patent used by the UHF RFID pool is one that is ‘necessarily essential to the standard (ie, inevitably infringed by compliance with the [Gen-2] standard)’ or one that is ‘essential to the standard as a practical matter because there are no economically viable substitutes . . . ie, not reading on the standard itself but nonetheless required to manufacture a competitive product compliant with the standard, due to production or design costs, consumer preferences, or other reasons . . .’\footnote{14}

The 3G Patent Platform for Third Generation Mobile Communication Systems presented different antitrust issues for patent pools.\footnote{15} The 3G Patent Platform refers to the International Mobile Telephony-2000 project, which includes five different radio interface technologies (‘platforms’) for use in third generation digital wireless telecommunications systems for voice and data that are approved by the International Telecommunication Union. The 3G Patent Platform is actually five independent platform companies, each with its own licensing administrator, although the 3G Patent Platform will provide some administrative and educational services that are relevant for all of the five platforms.

The DOJ Business Review Letter for the 3G Patent Platform did not object to an arrangement in which each Platform Company limited patents to those that are technically essential to comply with the individual platform standard. Each Platform Company has sole responsibility for all licensing and pricing with respect to its own 3G technology. In this respect, the 3G Patent Platform is similar to five separate patent pools.

By coordinating the activities of five potentially competitive platforms, the 3G Patent Platform raises the possibility of higher royalties than would occur if all negotiations were done with fully separate entities for each platform. The 3G Patent

\footnote{12} Business Review Letter (Jun. 10, 1999), n. 7 above (‘So long as the patent expert applies this criterion scrupulously and independently, it is reasonable to expect that the Portfolio will combine only complementary patent rights, and not limit competition between them and other patent rights for purposes of the licensed applications.’).


\footnote{14} Ibid.

Platform restricts the scope for coordination in the setting of royalties by limiting the role of the larger pool to suggesting standardized license agreements without price terms for use at the discretion of the separate Platform Companies, but the arrangement does not entirely eliminate the possibility of coordinated royalties. The DOJ concluded that: ‘We believe that the Platform arrangements have been reasonably structured to preserve the efficiency-enhancing integration of the identification and evaluation functions, without foreclosing competition in the critical aspects of the licensing and royalty-setting process by ensuring that these functions remain separate for each of the five technologies.’

These DOJ Business Review Letters and the conclusions in the IP Report set out the policies of the enforcement agencies with respect to the types of intellectual property that can be licensed by a patent pool without raising antitrust concerns. Similar antitrust concerns apply in some circumstances to cross-licensing arrangements. The statement of enforcement policies is a considerable advance relative to the history of antitrust decisions for patent pools, which I briefly review in Part II. Part III considers the theoretical underpinnings of the benefits and competitive risks of including different types of intellectual property in patent pools.

Part IV makes the argument that under some circumstances the risk of excluding patents from a pool that may prove to be essential is greater than the risk of including patents that are not essential. Over-inclusion does not harm competition provided that: (i) the pool includes at least one valid essential patent; and (ii) licensees are free to license the intellectual property they need to produce a product or implement a standard through independent negotiations with other rights holders. The risk of under-inclusion is that a pool may exclude a patent that is essential to implement a standard. Inclusion of patents in a pool that are not essential to produce a product or implement a standard does not harm competition if the pool contains at least one valid and essential patent and patentees are free to engage in independent licensing negotiations. A potentially greater concern is that owners of essential patents may be unwilling to join a patent pool that includes many non-essential patents. The presence of these non-essential patents would dilute their licensing revenues if royalty income were allocated in proportion to patents owned. In that case, owners of essential patents may prefer to license their patents independently, which would negate the pro-competitive benefits from the pool.

II. A Brief History of Patent Pools

Antitrust law has had a tortured history of enforcement decisions for patent pooling arrangements. In the early years of patents and antitrust, the view was that patentees

16 Ibid.
17 Cross-licensing arrangements differ from patent pools in that the former do not provide for centralized licensing of the patents involved in the cross-license.
had unbridled rights to assign, exchange, or combine their intellectual property rights. Courts soon recognized that patents were not exempt from the antitrust laws.

However, it was not until recently that the courts' analyses of patent pooling and cross-licensing arrangements approached the intellectual rigor expressed in advisory opinions by the DOJ and FTC, as summarized in their IP Report.\(^\text{18}\)


An early Supreme Court case on patent pooling is *E. Bement & Sons v. National Harrow Co.*, decided in 1902.\(^\text{19}\) After suing each other for patent infringement, six different firms assigned 85 patents dealing with float spring tooth harrows to National Harrow.\(^\text{20}\) Bement was a licensee. The pool grew to 22 firms accounting for over 90% of all manufacturing and sales of float spring tooth harrows in the United States.\(^\text{21}\) Each firm was required to adhere to uniform price schedules for the sale of all products manufactured under the National Harrow license. The pool set uniform license terms that fixed prices for licensed products, required that licensees make or sell only the licensed products, and obligated licensees not to challenge the patents and to defend the patents if challenged by others.\(^\text{22}\)

The Supreme Court did not engage in a detailed analysis of the antitrust implications of the National Harrow pool. In particular, the Court did not inquire as to whether the patents included in the pool were actually blocking or perhaps covered technologies that were substitutes for each other.\(^\text{23}\) The Court held that the licensing terms were lawful, even though they fixed prices for licensed products and prohibited the manufacture or sale of unlicensed products. The Court reasoned that the pool was legal because:

\[\text{[T]he general rule is absolute freedom in the use or sale of rights under the patent laws of the United States. The very object of these laws is monopoly, and the rule is, with few exceptions, that any conditions which are not in their very nature illegal with regard to this kind of property, imposed by the patentee and agreed to by the licensee for the right to manufacture or use or sell the article, will be upheld by the courts. The fact that the conditions in the contracts keep up the monopoly or fix prices does not render them illegal.}\]^\(^\text{24}\)

\(^\text{18}\) See IP Report, n. 1 above.

\(^\text{19}\) 186 U.S. 70 (1902).

\(^\text{20}\) A harrow is an agricultural device for spreading crop residue on fields, usually before planting.


\(^\text{22}\) See ibid.

\(^\text{23}\) See *National Harrow*, 186 U.S. 70. Blocking patents have overlapping claims such that the invention claimed in one patent cannot be practiced without infringing the claims of the other patent, and vice versa.

\(^\text{24}\) 186 U.S. at 91.
The themes expressed by the Court in *National Harrow* are that: (1) patent laws trump antitrust laws; (2) pooling arrangements confer benefits by avoiding costly litigation over patent scope and validity; and (3) licensing terms that fix prices are not unlawful because patentees have the right to specify the prices at which their products are sold. The potential harm to competition from arrangements such as the *National Harrow* pool is clear. Fortunately, the Supreme Court did not adhere for long to the principles enunciated in *National Harrow*, although the Court continued to struggle with the concept of whether patents are substitutes or complements for each other.


Ten years after the *National Harrow* decision, the Supreme Court did an about-face in *Standard Sanitary Manufacturing v. U.S.*, upholding the breakup of a joint licensing arrangement for patents relating to an enameling process for sanitary ironware.²⁵ The licensing arrangements established a standard royalty for the licensed patents and prohibited the marketing of products of inferior quality manufactured using the licensed patents.

As in *National Harrow*, the Court did not explicitly consider whether the patents involved in the Standard Sanitary licensing arrangement were blocking, complementary, or substitutes for each other. Interestingly, the Court did suggest that the licensing arrangement eliminated blocking positions. In this respect the arrangement was potentially pro-competitive, yet this fact received essentially no weight in the Court’s antitrust evaluation.²⁶

**C. Efficiencies defense for patent pools: complementary patent rights**

Two decades later, the Supreme Court specifically addressed the issue of blocking patents in its review of *Standard Oil Co. v. United States*.²⁷ In this case the Court reversed a district court finding that Standard Oil of Indiana and others had created an illegal patent pool to combine patents related to gasoline cracking, a key process in the refining of crude oil into gasoline.²⁸ The analysis focused on whether the cross-licensing of blocking patents violated the antitrust laws.

The Court found that the licensing agreement in *Standard Oil* contained none of the terms held to violate the antitrust laws in *Standard Sanitary*. In particular, the licensing terms did not restrict the freedom of the defendants individually to issue licenses under their own patents and did not impose any restrictions upon the

²⁵ 226 U.S. 20 (1912).
²⁶ Ibid.
²⁸ 283 U.S. 163.
quantity of gasoline to be produced or upon the price, terms, or conditions of sale, or upon the territory in which sales might be made. ‘The only restraint thus charged is that necessarily arising out of the making and effect of the provisions for cross-licensing and for division of royalties.’

The Court concluded that none of the patents involved in the pool was fundamental, but that each of the defendants had developed a cracking technology that arguably infringed other defendants’ patents. Most of the patents in the Standard Oil pooling arrangement were improvements upon other inventions. The basic inventions could block the use of the improvements, but the improvements did not prevent the use of the basic inventions.

D. Fixing royalties v. fixing product prices

The Supreme Court emphasized the benefits of cross-licensing of complementary patents in the Standard Oil pooling arrangement. Yet the Court did not perform a detailed evaluation of whether the patents in the pool were actually blocking or complementary, or whether some might have been substitutes for each other. Instead, the Court took comfort in the observation that the pool did not fix product prices or limit the use of non-licensed technologies.

The Supreme Court returned to the issue of cross-licenses involving patents that potentially block improvements in United States v. Line Material Co. The Southern States Equipment Corporation held a patent on a particular type of electric circuit protection device. Line Material held a patent that improved on the patented technology owned by Southern. The Court recognized that a cross-license between Southern and Line Material would be necessary for either company to exploit the technology inherent in both patents.

Southern and Line Material entered into a cross-licensing agreement whereby Southern made Line Material the exclusive licensor of Southern’s patent. The agreement gave Line Material the power to fix prices for devices that embodied both patents. The Court held that this power to fix prices under both patents was anticompetitive. The Court emphasized that cross-licensing to promote efficient production is not by itself unlawful, thereby confirming the conclusion that it reached in the Standard Oil case. ‘There is nothing unlawful in the requirement that a licensee should pay a royalty to compensate the patentee for the invention and the use of the patent. The unlawful element is the use of the control that such cross-licensing gives to fix prices.’

---

29 283 U.S. at 170.
30 283 U.S. at 174–76.
31 283 U.S. at 176–77.
32 333 U.S. 287 (1948).
33 333 U.S. at 315.
The key distinction between *Standard Oil* and *Line Material* is that the cross-licensing arrangements in the former did not explicitly fix prices for gasoline made with the licensed technology, although the *Standard Oil* cross-licenses did specify royalties, which of course have an impact on product prices.

### E. General observations on legal opinions for patent pools

A key issue emphasized in the DOJ/FTC IP Report is whether a patent pool or cross-licensing arrangement involves patents that are substitutes or complements for each other. \(^{34}\) Two patents are complements if they are each essential to practice a technology, or if a license to use one of the patents increases the value of a license to use the other patent. The first case involves two-way blocking patents, because each patent can block the use of the other patent. The second case involves one-way blocking patents, because one of the patents can block an improvement made possible using the other patent.

In most of the 20 antitrust decisions involving patent pools that I reviewed, spanning the period from 1902, when *National Harrow* was decided, to the present, the question whether patents were substitutes or complements was not the major determinant of enforcement outcomes. \(^{35}\) Table 11.1 below summarizes the enforcement outcomes. Of the 20 cases, 8 clearly involved patents that were complementary in that they were either essential to practice a technology or were in a one-way blocking relationship. In the other 12 cases, the patents either clearly were not complements or the court’s decision gave no indication of evidence that the patents were substitutes or complements. In the first group of 8 cases, 38% of the arrangements were held to be unlawful. In the second group of 12 cases, 42% were held to be unlawful. The difference is not statistically significant.

Anticompetitive effects from patent pools and cross-licensing arrangement can stem from several sources. These include the competitive relationships of the patents in the arrangement; the competitive relationships of products sold by participants in the arrangement; the presence of vertical restrictions in licensing terms or other ancillary restraints that affect competition related to the patented products.

<table>
<thead>
<tr>
<th>Did the arrangement include patents that were two-way or one-way blocking?</th>
<th>Number of arrangements</th>
<th>Number held to be anticompetitive</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>8</td>
<td>3</td>
<td>38%</td>
</tr>
<tr>
<td>No (or not known)</td>
<td>12</td>
<td>5</td>
<td>42%</td>
</tr>
</tbody>
</table>

\(^{34}\) See IP Report, n. 1 above.

\(^{35}\) See Gilbert, n. 2 above.
or processes; and agreements not to challenge the patents or to cooperate in the
defense of the patents.

For 20 pooling and cross-licensing arrangements, the most important determi-
nant of a verdict that the arrangement was unlawful was the presence of licensing
terms that fixed prices or divided markets for downstream products. Clearly, a
patent pool might be no more than a subterfuge for competitors to join together and
fix product prices or divide markets. This presumes, however, that members in the
pool could compete with each other if they did not have licenses to the patents held
by the other pool members. If patents are valid and blocking, such competition is
not possible, and in this case the agreements among the pool members does not
prevent competition that could occur in the absence of the pool. Analysis of the
antitrust risk posed by a patent pooling arrangement should look first to whether
the patents involved in the arrangement are substitutes or complements. While
courts have become more aware over time of the importance of addressing whether
patents are substitutes or complements, this is an exercise that has often been
overlooked.

### III. Theoretical Considerations

Patent pools are pro-competitive when they combine valid intellectual property
rights that are complementary and do not impose licensing terms that restrict
downstream competition. IP rights are complementary if they are essential to use a
technology or if they increase the value of other intellectual property.

#### A. Essential patents

Economic theory supports the conclusion that pooling of valid and essential patents
is pro-competitive absent licensing terms that restrict downstream competition. Suppose \( N \) patents are essential to practice a technology. Demand for the technol-
gy depends on its price, which in turn depends on the total royalties charged for
use of all of the patents. Furthermore, suppose royalties are charged on a per-unit
basis based on sales (running royalties). Let \( R \) be the total royalty for all \( N \) patents,
and \( r_i \) the royalty rate for the \( i \)th patent. The total royalty required to practice a
technology for which the \( N \) patents are essential is the sum of all of the individual
royalties:

\[
R = \sum_{i=1}^{N} r_i
\]

As an illustration, suppose that the demand for the licensed
technology is a linear function of the total royalty, \( D(R) = A - bR \), and licensing
incurs no marginal cost. The total revenue that a pool would earn by licensing the \( N \)

\[\text{See ibid.}\]
essential technologies is $RD(R)$. The total royalty that would maximize the pool’s revenue is the monopoly price, $R^* = \frac{A}{2b}$.\(^{37}\)

Acting independently, each of the $N$ patentees would choose a royalty $r_i$ to maximize its own licensing revenue: $r_iD(R)$. Let $R_{-i}$ be the sum of the royalties charged by all of the $N-1$ other technology rights holders other than $i$. The demand for a license from the $i$th patentee is the residual demand $A - b(R_{-i} + r_i)$, which takes into account the royalties charged by the other rights holders. If the licensor believes that its royalty rate has no effect on the royalties charged by the $N-1$ other rights holders, then its profit-maximizing royalty is $r^*_i = \frac{A - bR_{-i}}{2b}$. This is the monopoly price for its residual demand. The patentees are identical, so the profit-maximizing royalty $r^*_i$ is the same for all patentees ($r^*_i$) and $R_{-i} = (N-1)r^*$. Consequently, $r^* = \frac{A}{(N+1)b}$, and the total royalty required to license all of the $N$ complementary technologies is $R^* = Nr^* = \frac{NA}{(N+1)b}$.\(^{38}\)

Observe that for $N > 1$, the total royalty with independent licensing exceeds the pooled royalty, and the difference increases with $N$. As $N$ becomes large, the total royalty approaches the value that eliminates all demand for the technology. These results are in part a consequence of the assumption of per-unit running royalties.\(^{39}\) However, even with fixed fees for the licensed technologies, independent licensing can cause coordination difficulties that can lead to failures to obtain all of the licenses necessary to use the technology.

This example is a special case of the ‘double monopoly’ or ‘double marginalization’ problem encountered when manufacturers and distributors each add a markup to a good. A manufacturer and a distributor are in a complementary relationship that is similar to the relationship of licensors of essential patents. Acting independently, each patentee ignores the effect of its royalty on the licensing revenues of other patentees. However, the demand for licenses depends on the total royalties charged by all the licensees, and so the failure to coordinate royalty rates for essential and valid patents can result in total royalties that exceed the royalty that a pool would charge to maximize its licensing revenue. In addition, independent licensing of essential patents can delay the adoption of a technology and increase the transaction costs of negotiating licenses.


\(^{38}\) See Gilbert, n. 2 above. Shapiro derives a similar result using a different demand function. See Shapiro, n. 2 above. Lerner and Tirole generalize the result to examine the competitive effects of pooling patents that are imperfect complements. See Joshua Lerner & Jean Tirole, Efficient Patent Pools, 94 Am. Econ. Rev. 691 (2004).

\(^{39}\) A similar result would obtain with ad valorem royalties that are based on licensing revenues.
B. Improvement patents

Patent pools may include patents that are not essential to use a technology but enhance the technology’s value, whether by making it easier or more efficient to employ the technology or by enabling desired features. Common usage assigns the term ‘improvement’ to these patents, although they can add value in ways other than enabling a higher quality product. For example, a patented pin configuration can add value to a standard for a computer memory device even if the pin configuration does not improve the performance of the device relative to other configurations. Essential patents can block the use of these improvement patents, but the improvement patents do not block the use of the technology in a more basic form. This is a one-way blocking relationship. The essential patent is sometimes called the dominant patent.

As with independent licensing of essential patents, the ‘double monopoly’ problem also can emerge with independent licensing of improvement patents. In Line Material, if Southern had licensed its dominant patent at a uniform per-unit royalty, Line Material would have added its own margin to the price of the final product.40 As a consequence, the price of the final product could have exceeded the profit-maximizing price with a jointly determined royalty. Cooperative determination of royalties was a way for Southern and Line Material to avoid the double monopoly problem and could have resulted in lower prices for the final product.

The following examples illustrate conditions under which joint royalty setting with an improvement patent may or may not confer consumer benefits. Suppose first that all consumers are willing to pay for one unit of the good a price equal to the good’s value, which is known to all consumers. Firm 1 (eg, Southern) has a dominant patent that it can exploit to produce a product with value $V_1$. Firm 2 (eg, Line Material) has a subservient patent that, when used with Firm 1’s technology, can produce a product with a value (eg, quality) $V_2$ that is larger than $V_1$ for all consumers. I assume that there are no costs of producing or selling the products other than any royalties required to license patent rights.

If the firms cooperate, Firm 1 would allow Firm 2 to use its intellectual property to make the higher quality product. The two firms could license both patents jointly at the royalty $V_2$, or Firm 2, after having obtained a license from Firm 1, could issue a license for the higher quality product at the royalty $V_2$. This way the firms maximize profit, consumers purchase the higher quality good, and the firms can choose how to divide the spoils. In terms of total economic welfare (profits plus consumer surplus), this is the best outcome in this market, although the benefits go entirely to the firms. Consumers earn no surplus because the price is equal to their maximum willingness to pay for the good.41

40 333 U.S. 287.
41 Firms can perfectly price discriminate in this example because all consumers have the same willingness to pay for the good. Price discrimination would be more difficult, and consumers likely would enjoy some surplus, if they differed in their demands for the good.
If the firms do not cooperate, Firm 1 could choose not to license to Firm 2. With no licensing, Firm 2 cannot compete because its product would infringe Firm 1’s patent. In this case, only Firm 1 can offer a product, which it can sell at its monopoly price $V_1$. This outcome is inferior to the cooperative outcome because the higher quality good is not produced. The firms earn lower total profit, and consumers obtain no surplus.

Alternatively, Firm 1 could license its dominant patent to Firm 2 and commit not to compete with Firm 2 by withdrawing from the market. In this case, only Firm 2 offers a product, which it would sell at the monopoly price $V_2$. Firm 1 can extract the monopoly revenue by charging a royalty $R = V_2$. The outcome is the same as in the cooperative case. There is no double-marginalization in this licensing case, because I have assumed that demand for the final product is inelastic for all prices up to the product’s value. Firm 1 can capture all of the value with its royalty and Firm 2 cannot profitably add an additional margin to the product price. With elastic demand, if Firm 1 charged a per-unit royalty, Firm 2 would add a margin to the royalty and the double monopoly problem would emerge.

The outcome with licensing is identical to the outcome with pooling of the patents when consumer demand is inelastic. With elastic demand, there would be an additional margin in the no-pooling case if Firm 1 charges a per-unit royalty, which would increase the price and lower profits. Pooling is likely pro-competitive when demand is elastic and the pool includes both essential and improvement patents, as well as for the case in which all patents are essential, assuming that the patents are valid and restrictive license terms do not limit downstream competition. Stated differently, patent pools should not be limited to patents that are technically essential to make a product or implement a standard. Improvement patents are essential to obtain the benefits of some features of a technology or standard, but they are not technically essential to make products that lack those features or to implement a standard in its most basic form. Patent pools that include improvement patents also have economic benefits.

IV. Does it Matter if a Pool Includes Substitute or Weak Patents?

The antitrust agencies have focused on whether a patent pool is limited to essential patents as a central issue for a pool’s competitive effects. The IP Report notes that a pool that includes patents covering substitute technologies could result in higher prices by eliminating competition between alternative technologies and by facilitating collusion through the exchange of competitively sensitive information.42

While ‘essentiality’ may be accepted as a desirable characteristic of patents that are included in a pool, the term ‘essential patent’ admits different definitions, and the patents that are technically essential for a standard can change if the standard

42 See IP Report, n. 1 above, at 67.
evolves over time. If additional features are added to a standard, does the definition of an essential patent expand to include patents that are necessary to implement these new features? The MPEG-2 patent pool included 27 essential patents when the DOJ issued its business review letter in June 1997. In 2009, the portfolio license for the MPEG-2 technology provided access to more than 885 patents. The agencies recognize that it is difficult to assure that a pool includes only essential patents.

The use of an independent patent expert to assess essentiality provides some comfort, but essentiality is often difficult to determine even for an unbiased expert. Fortunately, the Agencies may relax their scrutiny of essentiality without compromising the economic benefits of patent pools. As noted above, patent pools that include complementary patents have economic benefits even if the patents are not technically essential to make or use some products or to implement the core features of a standard. Furthermore, the risk to competition presented by a patent pool that includes substitute patents is not large if the pool includes at least one valid essential patent. Moreover, this risk is less than the risk to competition from inadvertently excluding an essential patent. It is important to keep this tradeoff in mind given that it can be difficult to determine ex ante whether an individual patent is a substitute or a complement for other patents in the pool.

A dramatic example of the costs of excluding essential technology from a pool is the recent litigation between Lucent-Alcatel and Microsoft over patent rights to MP3 technology, which is used to transmit compressed audio files on the Internet and to store them on personal computers and portable devices.

There is no single pool with all of the patents that are essential to the MP3 digital audio encoding format technology. Ownership of MP3 patent rights is fractured among many different rights holders, including Fraunhofer IIS, Thomson, Sisvel (and its US subsidiary, Audio MPEG), Texas MP3 Technologies, and Lucent-Alcatel. Microsoft asserted its belief that it had obtained the necessary intellectual property rights to practice the MP3 technology when it negotiated a package license from Fraunhofer IIS for $16m. Lucent-Alcatel disagreed, claiming that Microsoft and its licensees infringed two patents necessary to use MP3 digital audio owned by Lucent-Alcatel. Lucent-Alcatel initially won a $1.5 billion judgment. The judgment

---

43 See Business Review Letter (June 26, 1997), n. 3 above.
45 See IP Report, n. 1 above, at 4 ("In many cases, patents in a pool are not pure complements or pure substitutes, but display characteristics of both.").
46 See Lucent Tech., Inc. v. Gateway, Inc., 543 F.3d 710 (Fed. Cir. 2008). MP3, which stands for MPEG-1 audio layer 3, is an audio compression technology. It is an enhancement to MPEG-1, and distinct from MPEG-3, which was developed for the compression of high definition television signals. See JERRY Whitaker,DTV Handbook: The Revolution in Digital Video 168 2001).
was subsequently overturned for non-infringement, but litigation continues between the companies on other patents.\footnote{See 509 F. Supp. 2d at 938 (granting Microsoft’s motion for a new trial); \textit{aff’d}, 543 F.3d 710. The Federal Circuit dismissed Lucent’s infringement claim on another patent on procedural grounds. See 543 F.3d 722.}

The Lucent-Alcatel MP3 case is an illustration of the risks of under-inclusion of essential patents in a patent pool. Whether the initial $1.5 billion judgment represents a patent ambush, or merely a stacking of royalties on top of the $16m for the Fraunhofer license, the point is that the total cost of licensing MP3 technology likely would be lower if all of the necessary patents were available from a single pool.

\section*{A. Should patent pools include patents that are substitutes for each other?}

Antitrust enforcement agencies, in their business review letters and policy statements, have indicated that patent pools should exclude patents that are substitutes for each other. The following example illustrates that competition need not be harmed if a pool includes substitute patents, provided that the pool also includes at least one valid essential patent.\footnote{The examples in this section follow the analysis of package licensing in Richard Gilbert & Michael Katz, \textit{Should Good Patents Come in Small Packages? A Welfare Analysis of Intellectual Property Bundling}, 24 INT’L J. INDUS. ORG. 931 (2006).} Suppose there are three patents that potentially read on a standard for a new sound recognition system. Patent X is technically essential to practice the technology. The other two patents, Y1 and Y2, read on some desirable feature of the technology that is included in the standard. For example, patent X is necessary to recognize any sound, including spoken words, while patents Y1 and Y2 cover technologies that are useful to recognize different types of music. Either Y1 or Y2 is essential to implement the music recognition feature of the technology, but either one will accomplish the same function. In this respect, patents Y1 and Y2 are perfect substitutes.

According to the guidance in both the DOJ business review letters for patent pools and the IP Report, patent X should be included in a patent pool for sound recognition technology because it is essential to practice the standard. Patents Y1 are Y2 are not essential because each one has a perfect substitute (Y2 for Y1 and Y1 for Y2), and guidance suggests that these patents should be excluded from the pool.

Patent pools can harm competition if they involve patents that can be used to make or sell products or processes that are substitutes for each other. This, however, is not a reason to exclude patents Y1 and Y2 from a pool that also includes a valid patent X. While patents Y1 and Y2 are substitutes for each other, they cannot be used to make or sell substitute products. In this hypothetical, any speech recognition product, whether or not it offers music recognition, must have a license to use patent X. Furthermore, by assumption, patents Y1 and Y2 are both useless without patent X.
Taking this example further, suppose that all consumers are willing to pay $V_1$ for a product that offers basic speech recognition, and that $V_2 > V_1$ for a product that includes music recognition as well. Demand is inelastic up to a price of $V_1$ for the basic speech recognition product and inelastic up to a price of $V_2$ for the product with music recognition. Suppose also that there are no costs of producing or selling the products other than any royalties required to license patent rights, and the marginal cost incurred by a patentee to license its patent is zero. If the patent pool includes X, Y1 and Y2, the pool can charge a royalty $V_2$ for a package license that enables both speech and music recognition. If the pool includes only X and Y1, it can offer a package license for X and Y1 that enables speech and music recognition for a royalty $V_2$. This would eliminate demand for a license for Y2. The cost to consumers would be the same as if all patents were in the pool, but in the former case Y2 would share in the pool’s licensing revenues.50

Suppose the pool excludes both Y1 and Y2 because each is a substitute for the other. If competition between Y1 and Y2 eliminates royalties for these patents, then the pool can charge a royalty for X equal to $V_2$. A manufacturer of speech and music recognition systems could license patent X, pay a royalty $V_2$, and license either Y1 or Y2 at no charge or with a royalty that is close to zero. More generally, if the lesser of the royalties for a license to Y1 or Y2 is $R_Y$, then the pool can charge a royalty for X equal to $V_2 - R_Y$. Consumers would pay a total royalty $V_2$ for the intellectual property necessary to make or use a product with a value $V_2$.

In this example, with inelastic demand up to each consumer’s willingness to pay for the technology, consumer welfare and total profits do not depend on whether patent Y1 or Y2 or both are included in the patent pool.51 Whether the pool includes Y1 and/or Y2 does affect the distribution of licensing revenues for the different IP rights holders. If Y1 and Y2 are in the pool, they can share in the total pool revenues on a basis determined by the members of the pool. If either patent Y1 or Y2 is outside the pool, competition for licensing revenues between the pool and the excluded patent will dissipate royalties for the patent that is outside the pool.

If both patents Y1 and Y2 are outside the pool, competition between the two patents will dissipate the royalty that each can charge. If Y1 and Y2 compete imperfectly for licensing revenues, they can charge a royalty $R_Y > 0$, which would limit the royalty that the owner of X can charge to something less than the value $V_2$. Furthermore, as in the discussion of United States v. Line Material Co. and the more general analysis of patent pools with one-way blocking patents,52 if demand for the licensed technology depends on the total royalty, then independent

50 Including Y1 but not Y2 in the pool raises the concern that the patent license would foreclose demand for Y2, as the DOJ noted in its 3C DVD and 6C DVD business review letters. See above text accompanying nn. 5–12. However, the foreclosure would not adversely affect the price of speech and music recognition systems.

51 See Gilbert & Katz, n. 49 above.

52 See above text accompanying nn. 32–33.
determination of royalties for X and for either Y1 or Y2 can lead to double marginalization if the pool excludes both patents Y1 and Y2 and the owners of these patents compete imperfectly. Under these conditions, with price-elastic demand for speech and music recognition systems, consumers likely would be worse off with both patents Y1 and Y2 outside the pool, and total profits likely would be reduced as well.53

The competitive effects from including non-essential patents in a pool are an application of the general principle of one-monopoly rent. This principle ignores factors such as commitment and effects in external markets, which can limit its application to conduct such as tying arrangements. It is conceivable that bundling non-essential patents into a pool’s portfolio license would have anticompetitive effects similar to the effects of tying arrangements.54 However, as discussed in more detail below, there are significant competitive risks from excluding patents from a pool that prove to be economically necessary to implement a technology. Courts and antitrust enforcement authorities should weigh the risks of excluding patents from a pool against the risks of harm to competition from an over-inclusive patent portfolio.

B. Should patent pools weed out weak patents?

Antitrust agencies have expressed concerns about including patents of questionable validity in a patent pool. The IP Report notes that ‘[a]n invalid or unenforceable patent is not in a complementary relationship with other patents in the pool’ and that ‘. . . the presence of invalid patents in a pool could raise competitive concerns.’55 Clearly, a patent right is not essential for the manufacture, use, or sale of a technology if the patent is not valid. Patents, however, are not valid or invalid with a probability of one. In Lemley and Shapiro’s terminology, patents are ‘probabilistic’ rights.56 They are valid with some probability, and the relevant question is whether it is better to have a patent of uncertain validity inside or outside of the pool.

There are benefits to having essential patents in the pool even if it is not certain that the patents are valid, provided that one or more other patents in the pool are essential and valid. Furthermore, excluding from the pool a patent that turns out to be essential and valid can lead to excessive royalties. For example, suppose that a pool includes five patents, each of which covers some aspect of a technology that is essential to implement a standard to broadcast video signals. The market demand for products that employ the broadcast standard is \( Q = 100 - R \), where \( R \) is the product price.

53 While consumer welfare and total profits do not depend on whether Y1 or Y2 or both are included in the patent pool along with patent X in this example, a pool that includes only patents Y1 and Y2, but not patent X, would raise antitrust concerns by eliminating competition between substitute patents.


55 IP Report, n. 1 above, at 78.

56 Mark Lemley & Carl Shapiro, Probabilistic Patents, 19 J. ECON. PERSP. 75 (2005).
To further simplify the analysis, assume that there are no costs incurred to make or sell the products other than the royalties paid for necessary patent rights, and the marginal cost of licensing a patent is also zero.

If any one of the five patents in the pool is valid, the pool can charge the profit-maximizing royalty, $R^m = 50$. This holds even if other patents in the pool are invalid, since a user of the technology must have access to all of the essential technological elements. Suppose each patent has an independent probability $\frac{1}{2}$ of being invalid. The probability that at least one of the patents is valid is $0.97$. Thus the pool can charge $R^m = 50$ with probability $0.97$. With probability $0.03$ the pool’s patents are all invalid and the pool cannot earn any royalties.

Suppose that, consistent with a business review letter from the DOJ, the pool conducts a quadrennial review of its patent portfolio. The review shows that one of the patents has questionable validity, and according to its agreement with the DOJ, the pool administrator concludes that the patent should be excluded from the pool. Are consumers better off if the pool excludes the patent? It is possible and even likely that consumers would be better off if all of the five patents remain in the pool, compared to a situation in which one of the patents is excluded from the pool.

If a patent is excluded, the probability that at least one of the pool’s four remaining patents is valid falls from $0.97$ to $0.94$. This is a modest reduction. It is still almost certain that the pool can assert at least one patent that is essential to make, use, or sell the product, even if only four patents remain in the pool. Furthermore, the excluded patent may turn out to be essential, because the review may have falsely concluded that the patent is invalid. Suppose that the excluded patent turns out to be essential with probability $\frac{1}{2}$. Thus, with probability $0.5 \cdot 0.94 = 0.47$, a potential user of the video technology would have to license at least one valid and essential patent from the pool and another valid and essential patent from an external licensor. This is a classic double-marginalization problem. Applying the Cournot pricing formula, the price in this instance would increase to $R^c = 66.7$, an increase of about $33\%$ compared to the total profit-maximizing royalty when all five patents are in the pool and at least one patent is essential.

The expected consequences from excluding one of the patents from the pool depend on whether patents turn out to be valid. With all five patents in the pool, the probability that at least one of them is valid is $0.97$. Hence the royalty is $50$ with probability $0.97$ and zero with probability $0.03$. The expected royalty is $0.97 \cdot 50 = 48.50$. The expected royalty is more complicated with four patents in the pool and one patent outside the pool. The probability is $0.47 (= 0.5 \cdot 0.94)$ that at least one patent in the pool is valid and the patent outside the pool is valid. In this case the royalty

57 This number is one minus the probability that all the patents are invalid. The latter probability is $(0.5)^5 = .03$.

58 See Gilbert, n. 2 above, and the formula for the profit-maximizing royalty assuming Cournot pricing in the text. In this example, $N = 2$, $A = 100$ and $b = 1$. 
is 66.7. There is an equal probability that at least one patent in the pool is valid, but the patent outside the pool is not valid. In this case the pool’s profit-maximizing royalty is 50. With probability .06, none of the patents in the pool are valid. If the patent outside the pool is valid, its owner can charge the monopoly royalty of 50. The excluded patent is valid with probability \( \frac{1}{2} \), hence this event occurs with probability 0.3. If the patent outside the pool is also invalid, a technology user can obtain all of the required rights with zero royalty. This event occurs with equal probability .03.

Putting all these possible events together, the expected total royalty with four patents in the pool and one patent outside the pool is 
\[
0.47 \times 66.7 + 0.47 \times 50 + 0.03 \times 50 + 0.03 \times 0 = 56.4.
\]

This is about 16% higher than the expected royalty with all five patents in the pool. In expected value terms, consumers are better off when all five of the patents with uncertain validity are included in the pool compared to a situation in which one of the patents is relegated to a status outside the pool.59

This example demonstrates that it is potentially dangerous to consumer welfare and total profits to exclude complementary patents from a patent pool, provided that it is reasonably likely that the pool includes at least one valid essential patent. This point holds even if some of the patents are improvement patents that might be blocked by other essential patents in the pool.

### C. Risks of too many patents in the pool

The usual concern about including non-essential or weak patents in a pool is that the pool may coordinate royalties for patents that offer substitute ways to produce products. But including non-essential or invalid patents in the pool does not result in higher total royalties if there is at least one other valid patent in the pool that is essential to make the product. Under the conditions for which the theory of one-monopoly rent applies, the presence of a nonessential or invalid patent in the pool does not allow the pool to charge a higher royalty, nor does it foreclose an alternative technology from making a product, provided that at least one valid and essential patent remains in the pool, because the product cannot be manufactured, used, or sold without a license to the essential patent.60

Competitive concerns arise if patents can be used to make alternative products that are substitutes for each other. Suppose that patent X1 is essential to make a ‘red’ speech recognition system and patent X2 is essential to make a ‘blue’ speech recognition system. The red and blue systems are close substitutes for each other.

---

59 The argument against excluding the patent is weaker if the excluded patent is less likely to be valid. A similar calculation shows that if the probability that the excluded patent is valid is only \( \frac{1}{4} \), then the exclusion increases the expected royalty by about 8%.

60 I reiterate that, under some conditions, bundling invalid or non-essential patents into a pool’s portfolio license could have anticompetitive effects similar to the effects of tying arrangements.
Including X1 and X2 in the same pool presents a risk that the licensing administrator for the pool will set royalties for X1 and X2, recognizing that a low royalty for one of the patents will cannibalize revenues for the other patent. That is, the licensing administrator may act as a cartel and eliminate competition between technologies that are substitutes for each other.61

While inclusion of patents X1 and X2 in the same pool raises potential antitrust concerns, it does not necessarily lead to a conclusion that competition would be harmed. Suppose the owners of X1 and X2 could still negotiate independently with potential licensees. Then, each patent owner would have an incentive to maximize its own licensing royalties by offering a licensing deal that is more attractive than the deal offered by its competitor. Independent licensing is a reason why the DOJ did not object to the 3G Platform for Third Generation Mobile Communication Systems.62 The 3G Platform patent is essentially five different patent pools, each with a separate licensing administrator. Each platform is a potential substitute for the others, and therefore the pool raises concerns that the central administrator may act to avoid competition between the different platforms. Nonetheless, the pool would not eliminate competition if the licensing administrator for each platform acts independently to negotiate licenses for its platform.

Suppose that the 3G Platform Pool did not have independent licensing administrators for each of the five platforms. It is still possible that the individual patent owners would negotiate independent licensing arrangements that maximize their profits if the pool administrator does not discourage them from doing so. Competition could occur, although the risk of coordinated royalty setting would be present. Competition requires that IP rights holders independently seek arrangements to license intellectual property for the manufacture, use or sale of alternative products. If there is sufficient independent action to license IP rights, competition can occur even if patents that can be used for substitute technologies coexist in the pool.

Allowing pool members to license their patents individually provides a safety valve to protect against anticompetitive effects from including non-essential patents in the pool. If patents are substitutes, individual patentees can license their technologies to develop competing technologies or to promote the development of the pooled technology along alternative paths. Lerner and Tirole have shown that independent licensing would lower the total profits of a pool that includes substitute technologies.63 Thus, independent licensing can be a check against collusive royalties for substitute technologies.


62 See above text accompanying nn. 15–16.

63 See Lerner & Tirole, n. 38 above.
An additional risk from including patents in a pool that are not essential, either because they are substitutes or because they are likely to be invalid, is that these patents can dilute the incentive for the owner of an essential IP right to join the pool in the first place. Consider again the example in which patent X is essential to make a basic technology with value $V_1$ and patents Y1 and Y2 cover substitutes for an improvement to the basic technology, which has a higher value $V_2$. The previous section showed that with perfect competition, or with imperfect competition and inelastic demand, total royalties and profits are independent of whether the pool includes X, Y1 and Y2, or X and only Y1 (or only Y2), or only X. However, the distribution of profits, and in particular the profits earned by the owner of patent X, likely depend on whether the pool includes all three patents, or two patents, or only patent X. If the pool allocated royalty income in proportion to patent holdings, the owner of patent X would earn $V_2$ if it is the only patent in the pool and if there is perfect competition between Y1 and Y2. But if all three patents are in the pool and the pool allocates royalty income in proportion to the number of patents owned, then the owner of patent X would earn only $\frac{1}{3}V_2$. Similarly, if the pool includes X and Y1 (or Y2), and the owner of the excluded patent charges close to zero in order to find a willing licensee, then the owner of patent X would collect $\frac{2}{3}V_2$.

If a patent pool allocates royalty income to pool members in proportion to the number of patents owned, a firm with an essential patent may choose not to join a pool if the pool includes too many non-essential patents. Suppose the pool includes non-essential patents Y1 and Y2, each of which is owned by a different patentee. The owner of patent X can guarantee royalty income of at least $V_1$ if it refuses to join the pool. The owner can license patent X solely for use with the basic technology, and it can refuse to make the technology available for improvements. This would allow the owner of patent X to charge a royalty equal to $V_1$. If the owner of patent X joins the pool, and if royalties are allocated in proportion to the number of patents owned, it would earn only $\frac{1}{3}V_2$. In that case, the owner of patent X would be better off by staying outside the pool and refusing to make its technology available for improvements if $V_1 > \frac{1}{3}V_2$, which is equivalent to $V_1 > \frac{1}{2}(V_2 - V_1)$, that is, if the stand-alone value of patent X exceeds half of the incremental value of improvements. Note that the owner of patent X can earn more than $V_1$ if it licenses its technology for improvements but refuses to pool its patent with patents Y1 and Y2. This provides a further incentive for the owner of patent X to refuse to join a pool that includes both patents Y1 and Y2 if the pool distributes royalties in proportion to the number of patents owned by each of its members. However, as in the discussion of United States v. Line Material Co. and the more general analysis of improvement patents, independent licensing by a pool that includes patents Y1 and Y2

---

\[64\] See above text accompanying nn. 32–33.
(but not X), and by the owner of patent X outside the pool, would lead to double marginalization and result in higher royalties and lower total profits if there is downward sloping demand for the product made with the licensed patents and if competition between Y1 and Y2 does not eliminate royalties for these patents. Furthermore, independent licensing for X and Y1 or Y2 can incur other costs that can delay adoption of the patented technology or lower its value.

V. Conclusions

Patent pools have clear efficiency benefits when they include patents that are essential to make or use a product. By pooling essential patents and licensing them jointly, a pool avoids the double-marginalization that can occur with independent licensing, reduces the transaction and search costs associated with assembling rights to use essential intellectual property, and accelerates technology adoption.

Antitrust authorities have indicated concerns about pooling patents that are substitutes for each other, because joint licensing can eliminate competition between substitute technologies and lead to higher prices. This chapter argues that such concerns may be exaggerated if the pool includes at least one valid patent that is essential to make or use a product that employs the technology covered by the patents in the pool. If the pool includes at least one valid essential patent, under some conditions the inclusion in the pool of one or more additional patents that are substitutes for each other does not increase, and can decrease, the total profit-maximizing royalty that users have to pay for necessary technology. Bundling non-essential patents into a pool’s portfolio license can have anticompetitive effects similar to the effects of tying arrangements. Courts and antitrust enforcement authorities should weigh the risks of excluding patents from a pool against the risks of harm to competition from an over-inclusive patent portfolio.

A significant concern for competition policy is the risk that owners of essential patents may choose not to participate in the pool. If essential patents or desired improvement patents remain outside the pool, independent licensing of these patents can destroy the benefits of pooling for pricing, transaction costs, and technology adoption. Under-inclusive pooling can occur because the owner of an essential patent refuses to participate in a pool, or because the pool excludes a patent that it incorrectly concludes is non-essential. Either outcome can have adverse consequences for economic efficiency and consumer welfare. Antitrust authorities should give due consideration to policies that encourage owners of essential patents to join a pool, or that discourage owners of essential patents from participating in a pool’s portfolio license.
I. Benefits of Digital Technology for Research

The availability of scholarly material in digital format has created enormous efficiencies in the research process. As recently as ten years ago, scholarly research typically involved multiple trips to a library to search through card catalogues and print and electronic indexing services to identify and locate relevant material. The researcher’s work was limited by the library’s hours and, notwithstanding the availability of inter-library loan, the library’s holdings. Moreover, because print and electronic indexing services did not embody ‘full-text’ offerings of the contents but rather relied on the categorization skills of the indexer, they provided only a filtered view of the sources available to the researcher. Further, indices were only as helpful as the topic they covered and the time period of coverage. A library with extensive holdings is of little use if it does not house a specialized subject index or an index covering materials from earlier years. And it is not unusual for desired material to

---

* Nancy Kopans is General Counsel and Secretary of ITHAKA, a not-for-profit organization that helps the academic community use digital technologies to preserve the scholarly record and to advance research and teaching in sustainable ways. ITHAKA’s services include JSTOR, Portico, and Ithaka S+R, the strategy and research arm of ITHAKA.
be absent from the library stacks, either because it was checked out by another patron or because it was mis-shelved. Even if desired material can be found in a library, the material available in analog format may be in poor condition. Age and normal wear and tear makes paper vulnerable over time. In some materials, pages may be defaced or removed (either intentionally or unintentionally). For example, in many early journals, engraved illustrations—such as ones by Audubon—have been removed to be framed as art. Content in some print materials has been destroyed by users who have disagreed with the contents or wanted to thwart other researchers’ use. Regardless of the intent, the defacing and removal of portions of materials clearly reduces the value and usefulness of what remains for researchers. Further, because the materials are held in many disparate locations, without any central coordination it is difficult if not impossible to know how many comprehensive copies actually exist, or even if a single comprehensive copy does.

Digital technology has revolutionized the research process, enabling scholars to access valuable material at any hour from any location. A researcher with access to adequate bandwidth does not need to coordinate his or her schedule around limited library hours. Nor does a researcher need to limit him or herself to the holdings of his or her affiliated library. Material available from an endless range of sources can be identified, if not fully accessed. Material can be researched across content types—text, images, and audio—and across disciplines that may overlap in their coverage of certain topics, facilitating the serendipitous discovery of valuable material. Linking and increasingly sophisticated search technologies can direct a user to relevant source material the user might not have encountered relying on analog search strategies alone. The availability of material in digital media also furthers important preservation goals. It means that researchers no longer need to depend solely on a tangible record which can easily be damaged, lost, or otherwise made unavailable. Because digital versions of materials tend to come from centrally managed resources, their integrity can be maintained efficiently and reliably. In the case of rare material, digitization can enable continued access over time without compromising the integrity or the long-term preservation of the artifact copy.

Microform collections are subject to many of the same vulnerabilities as print versions. Libraries might not own the materials in microform, and those microform collections they do own similarly are subject to mis-shelving, theft, and damage.

Notably, in connection with master planning projects for libraries at Massachusetts Institute of Technology and Rice University, both universities found that the time when students primarily undertook their academic work was between 11:00 pm and 4:00 am, times when the libraries typically are closed. See Geoffrey T. Freeman, Library as Place: Changes in Learning Patterns, Collections, Technology, and Use, in LIBRARY AS PLACE: RETHINKING ROLES, RETHINKING SPACE (2005), available at <http://www.clir.org/pubs/reports/pub129/pub129.pdf>.
II. Aggregation Assists in Realizing the Benefits of Digital Technology

Although technology can improve the research process, it gives rise to unique challenges. Key among these challenges are improving mechanisms for facilitating the location and discovery of vast bodies of digital material; rights clearance to enable the reproduction, display, and distribution of material online; developing ways for users readily to ascertain the quality of various materials found in an online environment; and the implementation of business models that can ensure the ongoing viability and availability of valuable scholarly material. Aggregation of content is essential to the realization of the benefits made possible by digital technology and for overcoming its challenges. Not unlike the roles performed by traditional physical libraries and publishers, aggregators can, to give a few examples, facilitate rights clearance, discoverability of content, quality assurance, and preservation.

There are many aggregators in the field of scholarly materials and they exist in various sectors. Examples of some of the better known ones are:

- Commercial third-party aggregators: EBSCO, Ingenta, ProQuest;
- Non-profit third-party initiatives: ARTstor, HighWire Press, Internet Archive, JSTOR, Project Muse, and various university and library initiatives (such as the Million Book Project);
- Multiparty initiatives: Open Content Alliance;
- Commercial search engines undertaking scanning and indexing projects: Google Book Search.\(^{67}\)

A. Rights clearance

When publishers act as their own aggregators, they may have adequate rights to reproduce, distribute, and display intellectual content online without undertaking further clearance. Third-party aggregators face ongoing challenges in clearing rights. The process of obtaining permission can be costly, time consuming, and often unsuccessful. Traditional grants of rights imposing time and geographical limits on usage may not square well with the needs of online archives seeking perpetual use for preservation purposes or use on an international scale, consistent with the global reach of the Internet. In an important study of problems encountered in seeking to obtain permission to digitize books, Denise Troll Covey focused on efforts at Carnegie Mellon University to secure non-exclusive permission to digitize and provide open web access to scholarly books in connection with its Million

\(^{67}\) A description of these organizations and initiatives is provided in Annex 1 to this article.
The study concluded that ‘[t]he range of publisher responses and their requests for fees, restrictions, and caveats show a publishing industry that has in no way reached a consensus on how to respond to libraries’ growing desire to provide digital access to scholarly materials. Indeed, some publishers are not even aware of what rights they actually own.’ Among Covey’s findings were that 21% of the publishers, accounting for 19% of the titles in the sample, could not be located. In a final sample of 277 titles from 209 publishers, about half the publishers responded to the request letters, and somewhat more than one-fourth granted permission, enabling the institution to digitize and provide web access only to approximately one-fourth of the desired books.

New technologies are shifting the traditional roles of parties involved in the creation, dissemination, and long-term preservation of material. Whereas, in the analog world, rights holders relied on libraries and other entities for ongoing distribution and preservation of their materials, print-on-demand capabilities, the development of publisher websites, and publishers’ ability to maintain digital files of their works have altered the balance. Thus, in addition to the time-consuming nature of rights clearance, rights holders’ justifiable interests in controlling and/or monetizing their content or in drawing attention to their content with new business models can pose challenges for aggregators’ rights clearance processes.

Yet, to the extent they become established and trusted entities, third-party aggregators are positioned to develop the resources, business models, and brand awareness to undertake rights clearance and to serve other functions. These factors have resulted in success for organizations like the Open Content Alliance and JSTOR. It also should be noted that large quantities of certain types of materials may be less challenging to permission than others. Section 201(c) of the US Copyright Act has been construed to permit the copyright holder in collective works such as periodicals, to digitize — and permit third parties to digitize — the work without seeking permission from the freelancers who contributed the individual articles as long as the original selection, coordination, and arrangement of contributions is preserved.69 This provision allows organizations that digitize periodicals to seek permission from a central rights holder rather than from a multitude of underlying rights holders.

Diane Zimmerman has described in detail potential solutions to rights clearance well-suited for aggregators.70 As Zimmerman points out, and as has been the

---


69 See Greenberg v. National Geographic Society, 533 F.3d 1244 (11th Cir. 2008); Faulkner v. National Geographic Enterprises, 409 F.3d 26 (2nd Cir. 2005), cert. den. 126 S. Ct. 833 (2005). Notwithstanding § 201(c), contractual restrictions in agreements between publisher and freelance contributors can still impose limitations on digitization.

70 See Diane Leenheer Zimmerman, Can Our Culture be Saved? The Future of Digital Archiving, 91 Minn. L. Rev. 989 (2007); see also, Diane Leenheer Zimmerman, Cultural Preservation: Fear of Drowning in a Licensing Swamp, Chapter 2 in this volume.
experience of JSTOR, establishing an embargo period for material to be held by
third-party aggregators can be an attractive mechanism for encouraging rights
holders to permit their material to be made available from third parties. By requiring
a time lag before material can be made available to users via the third party
aggregators, the aggregators are not viewed as competing directly with rights hold-
ers’ sales of current materials. Notably, the value of this approach may vary across
disciplines. In certain disciplines, such as medicine, current material is far more
valuable than legacy material. In other fields, such as history and economics, the
diminishment in the value of legacy material may be less precipitous. Nonetheless,
as researchers become accustomed to the ease of digital access to legacy material,
they might place increasingly greater value on legacy material, which in turn could
impact rights holders’ willingness to work with third-party aggregators to distribute
their content or the terms of these arrangements.

Another approach suggested by Zimmerman is the implementation of compul-
sory licenses in connection with the preservation of material. Zimmerman herself
expresses well-founded hesitation over this model. As even fee-for-service dark
archiving is unlikely to result in significant financial return, even a modest compul-
sory license rate could undermine preservation efforts, which serve a public good,
and any funds received likely will need to be used for operational purposes.
Moreover, a dark archive is not likely to compete with rights holders’ revenue
streams. Recommendations for changes to Section 108 of the US Copyright Act to
permit certain qualifying libraries and archives to make copies of works in their
collections solely for preservation purposes is yet another mechanism by which the
important function of preservation can overcome impediments of rights clearance.71

B. Improving discoverability of content

For content to be locatable online, it cannot exist in a vacuum; it must be connected
to other resources or hosted in a location that positions it to be easily discoverable.
To the extent they house a broad range of integrated content, aggregators can facili-
tate researchers’ serendipitous discovery of relevant content in other disciplines
and introduce researchers to new content types. Material no longer need be segre-
gated by time, rights holders, discipline, or even content type. A researcher in the
field of anthropology arguably could discover journal articles, monographs, field
recordings, and videos that may traditionally have been associated with disciplines
other than anthropology.

Content also benefits from the addition of associated metadata—data identifying
the author, title, date, and other relevant information—and that metadata, or the
full-text of the content itself must be made accessible through search engines and
other resources. Aggregators can be effective in developing metadata to facilitate

the discovery of the content they hold. Metadata is a critical element for enabling the location and discovery of content and it is frequently made available to third-party resources precisely for this purpose. Where metadata is inadequate, even the highest quality material may not be findable online, unless the researcher has ready access to full-text searching of the material.

Further, aggregators are well positioned to collaborate with other businesses to facilitate broader discovery of and access to content. Examples include arrangements with commercial search engines for crawling and indexing purposes. Indeed, commercial search engines can facilitate resource location. Increasingly, material in the ‘deep Web’ is indexed by commercial search engines, and a search on Google or Google Scholar can provide helpful location and discovery of these resources. Notably, content indexed by commercial search engines but housed within an aggregated collection is likely to have a higher search ranking than content standing on its own and thus is more likely to be discovered by researchers.

Aggregators also are well-suited to work with third-party organizations to develop outbound and inbound linking capabilities for references and citations within content. CrossRef, which facilitates the linking and location of materials cited in scholarly works among a vast community of rights holders, is one such third-party organization. They also are positioned to implement ‘Open URL’ standards. This ensures that when copies of the same electronic article are provided by more than one institution or database, the link from citation to full text resolves to a copy that is accessible to the user.\(^72\) It thus enables researchers who have obtained an article citation to identify and gain immediate access to the version of the article licensed by the researcher’s affiliated institution.

### C. Quality assurance

A significant challenge for researchers using online tools is to discern the quality of material discovered, a situation made more difficult because the ranking systems used by commercial search engines typically point to what is popular, but not necessarily what is high quality. Aggregators can introduce standards and quality assurance mechanisms that render them trusted providers of important scholarly content. They can take on a role of vetting material, thereby collating material known to be of high quality. Consequently, the content’s association with the aggregator can become an indicator of the content’s quality. Thus, depending on the standing of the organization, aggregation can provide a ‘seal of approval’ for the works it contains. Further, the quality of material contained within a broader aggregation is enhanced by its integration with that material. Functionalities developed by aggregators, such as faceted search technology enabling users to define the

\(^72\) See <http://www.niso.org/kst/reports/standards?step=2&project_key=d5320409c5160be4697dc046613f71b9a773cd9e>, last reviewed on March 20, 2008.
parameters of material they are looking for, features such as ‘more like this’ capabilities, and other functionalities can make material more valuable than it is standing on its own.

D. Preservation

Aggregators—and particularly third-party aggregators—can assume important functions for the long-term preservation of scholarly material. Charging copyright owners with the responsibility of preserving their own material removes an important check on the quality of the preservation process. For example, it is questionable whether the party that created it will choose to preserve material that turns out to be controversial or embarrassing, or in times of budgetary constraints, that an organization whose principal focus of operations is the production of content will spend adequate resources on its preservation. The digital environment carries particular risks in connection with preservation owing to the dynamic nature of online material and the ease with which an organization can remove material from its website. One example is the removal in 2003, just before the beginning of the Iraq war, of an article from a database of a 1998 *Time* magazine article sidebar ‘Why We Didn’t Remove Saddam’ by George Bush [Sr.] and Brent Scowcroft. The explanation offered by *Time*/EBSCO was that the removal was a result of a clerical error in the indexing process; before that, the reason given was that the article was removed at the rights holders’ request. Some speculated that the reasons might instead have been political.73

Entrusting third-party aggregators with archiving responsibility improves the likelihood that all material, including controversial material, will be preserved. Moreover, third-party aggregators can develop internal systems and controls for efficient, centralized preservation. An important example of this service is Portico, with a mission to ensure the long-term preservation of ‘born-digital’ journal material.74 At present, Portico has ingested over six million articles from over 50 publishers. A key component of Portico’s efforts include normalizing the original source files of electronic journals it receives to a standardized archival format in order to ensure consistent preservation across a range of standards. In so doing, Portico ‘[r]elieves publishers of the obligation to convert their source files into archival formats and conduct future content or format migrations as technology changes’ and ‘[c]reates an opportunity to reduce or potentially eliminate current internal archiving activities and related costs.’75

---

74 See [http://www.portico.org].
75 See [http://www.portico.org/publishers/].
III. Challenges of Aggregation

Although developments in online technology and its uses may remedy some of these problems, centralized aggregation of materials also introduces many challenges along with its benefits.

A. Licensing

As previously discussed, one of the main difficulties of aggregation is obtaining licenses to incorporate content into online databases. Yet, in addition to the need to clear rights on the content side, aggregators must also think about how to license their own end products to the user and library communities. Libraries obtain licenses for countless resources, and the rights that are conveyed are not standardized, but instead are quite varied. The process of reviewing and negotiating license agreements can be extremely time-consuming. Further, libraries and their users have become accustomed to certain usage rights for content obtained in analog format. These rights traditionally have been set by the breadth and boundaries of copyright law and associated guidelines, which permit fair use, interlibrary loan, and rights associated with the First Sale Doctrine, among others. With the rise of digital collections, increasingly greater proportions of library holdings have shifted from being governed by copyright law to be governed by the terms of providers’ license agreements. Libraries and users would like licenses for aggregated digital content to mirror closely what would be allowable under copyright law for use of hard copies. They also believe that provisions in licenses for use of digital content should include a standard understanding of what constitutes the community of authorized users (ie, for academic institutions: currently-enrolled students, faculty, and staff as well as ‘walk-in’ users on the institution’s premises; for public libraries: staff and patrons); availability of remote access; maintenance of user privacy; and use of best practices in areas such as content identifiers, usage statistics, link resolution, accessibility for the visually impaired, authentication, and perpetual archiving solutions. Thus, among the challenges arising for aggregators in connection with archiving are the development of more consistent usage rights and obligations in licensing of their content. SERU is an example of an effort to establish consistent and uniform standards.\(^76\) In addition, providers of aggregated content face the challenge of their material being available on a global scale, and thus potentially subjecting them to causes of actions under international law.

---

B. Searching across multiple aggregators

So too, one faces the challenge of multiple aggregators and how to connect them. We have witnessed an explosion of distribution channels (blogs, Wikipedia, pre-print archives, institutional repositories) in addition to the types of aggregators noted above. How can one search across these resources? One method is federated search technology, ‘allowing search and retrieval to span multiple databases, sources, platforms, protocols, and vendors at once.’ Federated search technology is aimed at enabling libraries to offer portal environments for library users analogous to web-based searching provided by Google.77 While federated search technologies may facilitate discovery of resources across providers, the effectiveness of these services are limited by the lack of consistent metadata standards. In addition, as more and more material, including material held by aggregators, is indexed by commercial search engines, the need for federated search technology may diminish.78

C. Improving search results

Another challenge of aggregation is to improve the ability of researchers to cull relevant search results from the myriad search results delivered. A recognized feature of online searching even of aggregated resources is that a researcher can be bombarded by voluminous responses to search queries, and can find it difficult to distinguish relevant material from search when results throw up hundreds or thousands of query responses. In time, enhancements in search technology and the way in which the web is organized are expected to improve our ability to identify key content. Search technology of the future is expected to encompass not only key word and full text search capabilities, but also the use of technologies like computational linguistics, statistical analysis, and natural language process. The web likely will evolve to contain resources corresponding not just to media objects (such as web pages, images, audio clips, and so forth) as it currently does, but also to objects such as people, places, organizations, and events; it will utilize not just a single kind of relationship (the hyperlink) between resources, but many different kinds of relations between the different types of resources mentioned above. These will enhance the ability of users to locate relevant content and will improve efforts to synthesize and point users to resources used by like-minded scholars.

D. Access and authentication

Other considerations involve the technology that manages how access and authentication for the use of aggregated material is obtained. At present, most access to ‘gated’ content within the scholarly community is based on Internet Protocols or passwords, methods that allow for little flexibility in range of usage rights for content. Over time, access will shift from these methods to more granular methods based on ‘attributes.’ For example, Shibboleth is an initiative to develop an open, standards-based solution to organizations’ need to exchange information about their users in a secure manner that preserves users’ privacy.79 For example, through Shibboleth, resources can be identified for access by all active members of a campus community whereas other resources can be limited to students in a particular course.

E. Business models

Aggregators also face challenges in connection with maintaining sustainable business models. For example, open access80 and the desire for ready access to public domain material, while having the laudable aim of benefiting users (including particularly scholars not affiliated with major academic institutions or public libraries), can put pressure on business models that facilitate the long-term sustainability of important efforts. In evaluating the benefits of open access and broad access to public domain material, it is important to consider the ongoing costs of maintaining digital preservation sites. These expenses are significant and ongoing. They involve not only the cost of maintaining the website and its technology but also that of refreshing data—in some cases, re-digitizing content when specific needs arise—and increasing functionality to improve the usefulness of the resource. By way of analogy, traditional physical archives also house many public domain materials; but to keep them available, they must be able to pay for their real estate costs, their staff, and the equipment necessary for the long-term physical preservation of these materials. That money must be generated somehow. Although it may be reasonable in many cases to make material freely available after certain costs are recovered, it is important to recognize that if an entity is able to generate a revenue stream from its collection, whether it be public domain material or material still under copyright, it is more likely to stand on solid feet as a self-sustaining organization and will be better able to serve the public’s archiving needs.

80 ‘Open access’ describes an effort to make peer-reviewed scholarly articles freely available to end users via the World Wide Web, with only limited copyright and licensing restrictions.
Revenue generation also gives it the flexibility to experiment with adding additional material that might, for various reasons, be more costly to archive.

IV. Conclusion

Digital technology has transformed the research process by making a broad range of content types created over a vast expanse of time available any time of day from almost anywhere in the world. Aggregation is a critical component for realizing the benefits of digital technology. Aggregators can be well-positioned to undertake rights clearance, improve the discoverability of content, provide quality assurance, and facilitate preservation of material. At the same time, aggregators face numerous challenges. Rights clearance can be time consuming, aggregators need to be mindful of the ways in which their content is licensed to users and libraries, and the presence of multiple aggregators can create impediments for users seeking to search across aggregated resources. Notwithstanding these challenges, the value of aggregation is undeniable, and in time, developments in online technologies likely will address many of the more confounding challenges.

ANNEX 1

Descriptions of Aggregators

Commercial third-party aggregators
EBSCO: provides print and electronic journal subscription services, research database development and production, and online access to more than 150 databases and thousands of electronic journals. <http://www.ebsco.com>.

Ingenta: provides citations, brief abstracts, tables of contents, and some full text to articles in over 28,000 publications in a broad range of subject areas. <http://www.ingenta.com>.

ProQuest: a multidisciplinary resource with citations and abstracts for over 2,500 journals, magazines, and newspapers, many dating back to 1986 or earlier and covering all subject areas. Full text is available for about one-half of the publications and is searchable. <http://www.proquest.com>.

Non-profit third-party initiatives
ARTstor: a digital library of nearly one million images in the areas of art, architecture, the humanities, and social sciences with a set of tools to view, present, and manage images for research and pedagogical purposes. <http://www.artstor.org>. 


Internet Archive: an Internet library, with the purpose of offering permanent access for researchers, historians, and scholars to historical collections that exist in digital format. <http://www.archive.org>.

JSTOR: a collection of over 1,300 digitized, full-text searchable peer-reviewed academic journals, from over 500 publishers in more than 50 disciplines, in some instances dating back to the late seventeenth century. <http://www.jstor.org>.

Million Book Project: an international venture led by Carnegie Mellon University, and universities in China, India, and Egypt, which has completed the digitization of more than 1.5 million books, now available online. <http://www.archive.org/details/millionbooks>.

Project MUSE: a not-for-profit collaboration between participating publishers and the Milton S. Eisenhower Library at Johns Hopkins University, providing institutions with full-text online access to a comprehensive selection of humanities and social sciences journals. <http://muse.jhu.edu>.

**Multiparty initiative**

Open Content Alliance: a collaborative effort of a group of cultural, technology, nonprofit, and governmental organizations from around the world that is administered by the Internet Archive and helps build a permanent archive of multilingual digitized text and multimedia material. <http://www.opencontentalliance.org>.

**Commercial search engine scanning and indexing tools**

Google Book Search: a tool from Google that searches the full text of books that Google, in collaboration with library partners, scans, OCRs, and stores in its digital database. Google makes ‘snippets’ of these books available in response to user search queries.81 <http://books.google.com>.

---

81 Google’s activities in connection with Google Book Search gave rise to a law suit brought in October, 2005 by five publishers (McGraw-Hill, Pearson Education, the Penguin Group, Simon & Schuster, and John Wiley & Sons), who asserted that scanning a copyrighted work without permission does not constitute fair use and thus violated the publishers’ copyright interests. As a result of a proposed settlement in the case, and subject to final approval of the settlement, the Google Book Search project may expand to include online display and distribution of scanned versions of books, rather than display of only snippets.
The key rationale for granting intellectual property rights is practical: they can induce technological innovation and dissemination by at least partially privatizing associated social benefits. In the context of the basic, one-stage model of innovation this idea is persuasive. Innovation is work, frequently hard, risky, and frustrating, and its most valuable fruits, in the absence of intervention in the marketplace, tend to be available free to all, so none has an incentive to pay. Intellectual property rights give some degree of monopoly power to the holder, and the prospective value of this market power is necessary to encourage greater innovative effort, and more innovations. All this seems to be a classic application of the insights of Adam Smith’s *Wealth of Nations* regarding the importance of financial incentives for national economic progress.

When one invention builds on another, monopoly at any one stage negatively affects the incentives at others, so the logic becomes less obvious. However, if there is no invention without intellectual property, the latter can only improve the overall rate of innovation.

The history of technical progress in agriculture and manufacturing raises a number of questions about this argument. In agriculture, improved methods of planting, cultivating, fertilizing, and harvesting crops are easy to observe and to copy. Until the 1980s, there was no strong legal protection that could constrain use crop seeds for further breeding. Yet a multitude of studies has affirmed that the rate of technical progress in agriculture over the past century has been exceptionally high, in the United States and other developed countries.82

The fact that effective incremental mechanical innovation worked well in the agricultural sector in the nineteenth century with minimal private incentives attracted little public discussion. In contrast, after the merits of private incentives had captured the public imagination during the ‘Reagan revolution,’ an example in a new industry of effective collaborative innovation without high-powered incentives has generated much attention and *ex post* rationalization. Highly complex computer software, freely available for use by others, with openly available source

---

code, was generated by groups of programmers with little obvious prospects of reward beyond perhaps making the software a more efficient tool to meet their own needs. Moreover, it proved to be competitive, in some applications, with alternatives produced by for-profit private firms.

How could this be? One factor was itself an enabling innovation: a novel use of copyright licenses to ensure that the benefits of improvements by users would be available to all participants, and to prevent hijacking of the group efforts for private gain. This 'copyleft' strategy was well suited to protection of software code, and was implementable at low cost.

In agriculture, an opposing trend was set in motion in the 1980s. At the start of the decade, the twin revolutions in biotechnology and intellectual property were heralded in the United States by the grant of the Cohen-Boyer recombinant DNA patents and the Supreme Court’s *Diamond v. Chakrabarty* decision,83 affirming the patentability of new life forms.

Crucially, utility patents on life forms granted a monopoly on their use for breeding further new varieties. The first IPR specifically for plants had been introduced much earlier in the United States, by the 1930 Plant Patent Act, in the form of special ‘plant patents’ for new and distinct plant varieties.84 However, these restrict only asexual propagation (cloning) of plants. In this sense they are analogous to copyright restrictions on simple reproduction of creative works; they do not affect breeding of new varieties. The Plant Patent Act influenced the International Convention for the Protection of New Varieties of Plants (UPOV) established in 1961 by a group of Western European countries.85 This convention offers a model system for ‘plant breeders’ rights.’ In turn, the United States 1970 Plant Variety Protection Act, which covers production of sexually reproduced plant varieties but does not constrain their use in breeding, followed the UPOV model.86

The high-powered patent incentives for applying the new methods of plant transformation to freely available unpatented varieties elicited an impressive private-sector research response, by both startup firms and others already providing inputs for crop producers. This was one important driver of the rise in investment in agricultural research in the United States to match the public sector by 1995. Fuglie and coauthors found United States nominal private investment in agricultural biotechnology increased fourfold in the first 12 years after the *Diamond v. Chakrabarty* ruling.87 Mergers generated a set of life science corporations with plans to apply biotechnology for human health as well as for agriculture. King and Schimmelpfennig

---

show that by 2002 six firms, all with histories in the chemical or pharmaceutical sectors (including Monsanto, Dow, DuPont, BASF, Bayer, and Syngenta), controlled over 40% of the agricultural biotech patents owned by the private sector. They did not achieve this presence by superior research in this new area; the subsidiaries acquired by these firms through mergers and acquisitions responsible for 70% of their total patent stocks.

The focus of private sector research has been narrow but deep. Since 1987, over 55% of all field trials for genetically modified (GM) crops—an area of research dominated by private industries—have been for corn or soybean varieties. Global GM crop value and planted area is almost entirely in soybeans, cotton, corn, and canola. Monsanto has become the leading firm in generation and diffusion of agricultural biotechnology, focusing activities on corn, soybeans, and cotton incorporating herbicide tolerant ‘Roundup Ready’ technology and pest-resistant traits based on crystal proteins derived from samples of Bacillus thuringiensis (Bt). It is unlikely that public research institutions could have matched the efficiency of Monsanto and other leading firms in these activities.

Paradoxically, patent protection of biotechnology products used in further innovations led to an increase in mergers of firms that helped avoid the costs of transactions in intellectual property. An environment of fragmented and decentralized proprietary claims on information encouraged firms to consolidate in order to avoid transaction costs and exploit increasing returns to scale or market power. Notably, Marco and Rausser empirically show that in the plant biotechnology industry, the enforceability of a firm’s patent portfolio is a good predictor of participation in consolidation. The authors note also that a number of mergers, including the Monsanto-Calgene and Monsanto-DeKalb cases, occurred in the context of patent infringement litigation.

Life science firms with freedom to operate within their integrated corporate structure may see little advantage in licensing their patents to others. By blocking new firms from entering, market consolidation may suppress innovation that threatens the market for their own products, chemical or biological. There is mounting

---

evidence that multiple, mutually blocking intellectual property claims on inputs are hindering access of entrants to research tools, creating an anticommons problem in commercialization of agricultural research. The application of IPRs to plant components and processes imposes high transaction costs for researchers who must amass fragmented proprietary inputs to develop a single downstream innovation. In the field of agricultural biotechnology, economists have claimed that the United States Patent and Trademark Office has granted poorly specified patents that are too broad in scope.

Agricultural economists have long been concerned with the excessive power conferred by ‘blocking patents’ on locked-in but otherwise non-crucial genetic technologies. In the field of plant biotechnology in particular, where ownership of the genes, markers, or promoters incorporated in a single innovation was initially fragmented, upstream IPR-holders, unwilling to allow commercialization of varieties using their property, have in some instances foreclosed university development of new crops. Lock-in of technologies in the course of development and regulatory assessment magnifies the blocking rights associated with intellectual property, a point only slowly appreciated by lawyers and economists unfamiliar with the technology of transformation of life forms. Recently, the broader economics profession has become focused on blocking patents in the context of embedded software, which are in fact much simpler to ‘work around’ than those in agricultural biotechnology software. The Supreme Court appears to have acknowledged the blocking problem in its eBay decision, which somewhat constrained the threat of injunction that holders of blocking patents could use to extract high royalties from infringers.

---


95 See n. 8 above.

96 Examples include the cases of the GE tomato and herbicide-resistant strawberry at the University of California at Berkeley, a herbicide-tolerant barley, and a herbicide tolerant turf grass at the University of Michigan, see Wright, n. 8 above; Wright, Agricultural Innovation, n. 11 above; F.H. Erbish, Challenges of Plant Protection: How a Semi-public Agricultural Research Institution Protects its New Plant Varieties and Markets Them, in THE IMPACT ON RESEARCH AND DEVELOPMENT OF SUI GENERIS APPROACHES TO PLANT VARIETY PROTECTION OF RICE IN DEVELOPING COUNTRIES, Los Banos, The Philippines, 16–18 February 2000. International Rice Research Institute (2000). More recently, commercialization of transgenic hypoallergenic wheat technology has been blocked by a combination of patent protection and regulatory costs (P. Lemaux, personal communication). Generation of many further US examples is unlikely, given that independent implementation of new transgenic technologies by small startup firms or universities is widely regarded as economically infeasible. Importation of yellow Mexican beans into the United States, and pineapple research in Latin America, have been disrupted by threats of enforcement of dubious or invalid patents, see Wright & Pardey, n. 11 above.


Using principles reminiscent of the long history of farmer innovation and technology sharing in response to local conditions, researchers in the public and nonprofit fields have undertaken initiatives to ameliorate anticommons obstacles in agriculture. Two notable examples are the Biological Innovation for Open Society (BiOS) and Public-Sector Intellectual Property for Agriculture (PIPRA) initiatives.

The Australian nonprofit, Cambia, set out with a plan to license its IPRs (including the widely licensed GUS reporter gene patented by Cambia’s founder, Richard Jefferson) for biotechnology in agriculture to users in wealthy countries, while making them freely accessible in developing countries. Jefferson, inspired by the success of open source models in software, later formed the initiative entitled Biological Innovation for Open Society (BiOS), to license rights to use key research tools not subject to private-sector IPR, in exchange for a commitment to share any downstream improvements of the tools with all BiOS members.99 This initiative aimed to attack the key problem that freedom to operate in agricultural biotechnology was restricted by the holders of patents on essential research tools for transformation of plants and other organisms. Under the BiOS license for research tools, applications of the tools could however be protected by their inventors with patents. Despite its acronym, BiOS is not strictly open source, as it charges for membership on a sliding scale.

It appears that participants have been happy to access Cambia technologies provided by BiOS, but have not reciprocated by offering their own technologies under an open source license. A fundamental problem might be that any success achieved in open source software using copyright licenses to prevent appropriation of the core technology is difficult to replicate in a system that relies on patent protection and is subject to the time-consuming and very costly regulatory requirements imposed on agricultural biotechnology.100

In a 2003 Science article signed by 14 university presidents, chancellors, and foundation presidents, the authors highlight the negative effect of intellectual property rights on ‘freedom to operate’ in agricultural research.101 This paper outlined the rationale behind the formation of Public-Sector Intellectual Property for Agriculture (PIPRA), supported by the University of California and the Rockefeller and McKnight Foundations.

The original goal of PIPRA was to act as a clearinghouse for information about patenting and licensing of technologies originating in the public and nonprofit sector to facilitate their commercialization and adoption in less developed countries.

and by producers of ‘minor’ crops such as the grapes and almonds produced in California. The common problem of both target groups is that their markets are too small to attract much interest from the major agricultural biotechnology corporations. The intellectual property strategy, similar to that of Cambia, was to protect proprietary claims for commercial use in developed countries (consistent with federal policy expressed in the Bayh-Dole Act of 1980,102 and with the aims of university licensing objectives), while simultaneously providing access to users in less developed countries and producers of minor crops, and reserving for public and nonprofit institutions the rights to use the inventions for applications in developing countries.

Both Cambia and PIPRA spent years and substantial high-quality scientific resources on research on work-arounds for key blocking technologies for plant transformation. While these have produced notable scientific advances,103 and PIPRA is now making suites of DNA vectors and a marker-free plant transformation platform freely available for research and humanitarian use, neither Cambia nor PIPRA has thus far produced a technology package that has been widely accepted by modern breeders as an efficient, completely unencumbered alternative to existing patented transformation technologies. Together these attempts are strong evidence that the blocking effect of proprietary claims over key elements of plant transformation is powerful and indeed stronger than the management of PIPRA and Cambia, experienced in patenting of plants and plant transformation, had anticipated.104 The lack of attractive unencumbered alternative technology sets may be the main reason why efforts to encourage collaboration in open-source type ventures have not made more headway. However, had they already succeeded at this level, the regulatory hurdles would still have loomed large. As it is, one firm (Monsanto) currently accounts for almost two-thirds of all public and private plant biotech field trials in the United States.

Neither Cambia nor PIPRA now presents creation of unencumbered alternatives to key technologies as a major objective. Currently, both are emphasizing provision of easily accessible information to researchers in developing countries and nonprofit institutions regarding freedom to operate within the context of patenting as the dominant paradigm, very much in line with the accommodative stance of the conference upon which this book is based.

The lack of technology packages with freedom to operate has not greatly impacted the daily work of biologists. In fact, surveys show that researchers, public or private, routinely ignore any prior patents when they pursue their own research, even

103 See, eg, W. Broothaerts et al., Gene Transfer to Plants by Diverse Species of Bacteria, 433 NATURE 629 (2005).
104 A scientist later involved in research on such attempted work-arounds once characterized common use by academics of proprietary technologies as ‘sheer laziness.’
though many of them understand that there is no significant ‘research exemption.’\textsuperscript{105} For the vast majority of academic bench scientists who do not have direct responsibility for commercialization of innovations, the real problem with respect to freedom to operate is that their universities themselves insist (albeit with only partial success) on use of material transfer agreements (MTAs) governing exchanges of research materials between researchers, to protect prospective university revenues from licensing intellectual property rights and limit university liability.

In a recent survey of agricultural biologists at Land Grant Universities, over a third of the respondents reported delays in obtaining access to research tools in the five years preceding the survey, with a mean of two delays and a mean duration of over eight months.\textsuperscript{106} Researchers reported responding to hold-ups by substituting tools, sometimes of lower efficacy, and in some cases by abandoning the project altogether. Most respondents (including scientists who have obtained patents) view intellectual property protection as a net negative for progress of research in their fields.

In follow-up interviews, scientists attributed the majority of these delays to problems encountered by university administrators whom they view to be principally concerned with protecting financial claims. This view is validated by a survey of land-grant university administrators, who report that provision of research funds is the greatest advantage of university-industry relationships, but believe that scientists are most concerned with research enjoyment and satisfaction of curiosity.\textsuperscript{107}

Thus bench scientists perceive that the main effect of intellectual property protection on the effectiveness of their work acts through the intervention of administrators in the material transfer process, to protect their institutions’ financial interests. Moving from the project to the program level, there has been concern that increasing use of intellectual property protection by public agricultural research institutions might, quite apart from any effect on the day-to-day work of scientists, also have an influence on the projects they work on. For example, Just and Huffman note that while the Bayh-Dole Act led to a jump in university patenting (from less than 400 patents per year before Bayh-Dole to 1,100 per year by 1989 and over 3,000 per year in the 1990s), the Act may have reduced the pool of basic research supporting private research, shifting the focus of public research towards shorter-term incentives such as patents, at the expense of future public goods research.\textsuperscript{108}


\textsuperscript{106} Lei et al., n. 23 above.


\textsuperscript{108} R.E. Just & W.E. Huffman, *The Economics of Universities in a New Age of Funding Options*, revised and resubmitted to RESEARCH POLICY.
On the other hand a survey, which did not directly address the difficult basic/applied distinction, found that patents and publishing appear to be complementary and subject to scale economies.109

Indeed the greatest effect of patenting might be on the structure of research activity in the private sector. Here the domination of innovation by one firm, indicated by the field trial data, is associated with the disappearance of startup firms in the markets for agricultural biotechnology and commercial crop varieties. Costly and inflexible regulations and lack of freedom to operate have reduced the rate of innovation110 and reinforced the barriers to entry, and the relative contributions of each are difficult to discern.

The lack of success of open source biology in agriculture has its counterpart in the failure of open source drug discovery after a decade of work.111 Quite recently, some observers predicted open-source development of therapeutic molecules112 or wholesale independent development and trading of life forms,113 prospective outcomes which raised serious issues of safety and security. Recent proposals advocate enlisting open-source collaborators in those particular stages of integrated research strategies in which progress depends on the number of eyeballs engaged.114 This seems not too different from the long-standing practice of major corn seed firms, annually enlisting farmers as collaborators in conducting thousands and even million of crop variety trials on their land across the United States.

This brings us back to where we started. Just how new is open source, and how different from the Smithian paradigm? Many who cite Smith on the unique virtues of patents in eliciting innovative effort appear not to have read all the way through Chapter 1 of his Wealth of Nations.

A great part of machines … in those manufactures in which labor is most subdivided, were originally the inventions of common workmen, who, … employed in some every simple operation, naturally turned their thoughts towards finding out easier and readier methods of performing it. Whoever has been much accustomed to visit such manufactures must frequently have been shown very pretty machines, which were the inventions of such workmen in order to facilitate and quicken their particular part of the work.

We should stop regarding open source innovation as a new phenomenon of limited relevance beyond some software applications. We might start investigating

109 J.D. Foltz, B.L. Barham & K. Kim, Synergies or Tradeoffs in Life Sciences Research, 89(2) AMERICAN JOURNAL OF AGRICULTURAL ECONOMICS 353 (2007).
114 See Maurer, n. 29 above; Weber, n. 30 above.
whether and how it can continue to contribute, perhaps in modified form, in areas where innovation takes time, and requires costly regulation to protect public safety. In agricultural biotechnology the battle between a traditional model of public innovation and free exchange of crop breeding materials is largely over in the most lucrative markets in developed countries (corn, cotton, soy, and canola). Patents (many on publicly originated technologies) have combined with regulation and technology lock-in to prevent entry and support the domination of innovation and diffusion of new varieties in these markets by a small number of firms with one clear leader. In this context, initiatives to foster open-source alternatives for other crops are particularly attractive, but difficult to achieve in the current regulatory and intellectual property regimes.