Unbundling Copyright from Patents to Inform the Analysis of Notice Externalities and Monopoly

Stan J. Liebowitz

September 20, 2015

Preliminary

Abstract:
Copyright and patent law are often treated as two peas in a pod. Perhaps economists are guiltier of this than are lawyers, but I am not so certain. We all know that there are differences between these two fundamental components of intellectual property. When copyright critics use a particular brush, say monopoly, to tar one, it is the rare critic who is able to keep from claiming that the same brush can be used to tar the other. At least that has been my experience.

This is not to say that some important differences between copyright and patent have not been noted. But it is my belief that even when noted, the differences are not well understood in the literature even though these differences have been pointed out.

In this paper I wish to focus on what I believe to be two crucial difference between copyright and patents that, to my mind, requires very different analyses and conclusions. The newer difference is the treatment of notice costs, which I believe are of much more import in the case of patents then they are in the case of copyright. The older difference is in the claim that intellectual property creates a monopoly, which again, is possibly correct for patent but not at all correct for copyright. The reason that these two fundamental components of intellectual property discourse differ from one another is due to the fact that copyright, while preventing copying, does not prevent independent creation, whereas patent law prevents both copying and independent creation.

Because independent creation is allowed by copyright law, creators have a simple algorithm for avoiding infringement—do not copy other works. This renders notice costs irrelevant for the large class of creative works that do not copy prior copyrighted works. The real-world limitations on the copyright “property right” (such as duration and fair-use) raise information costs although it is still not clear that notice costs for copyright are particularly high relative to other types of properties, contrary to some claims in the literature. Patents, by not allowing independent creation, impose much higher notice costs on potential users of patented inventions than is the case for the costs imposed by copyright on users of creative works.

Copyrights and patents provide ownership rights to the creators of works and innovations, preventing others, without permission, from making copies or products that employ the creations that those rights protect. Analysts, and particularly critics of intellectual property often refer these ownership rights as “monopolies,” but this is an error in the case of copyright. Because copyright does not prevent independent creation, it does not create a monopoly. That does not mean that a copyrighted good cannot have a monopoly in a market. But if so, the monopoly is due to some other aspect of the work, not to copyright. Because patents prevent independent creation, they can create monopolies. A patent is considerably more than merely a property right over a particular innovation. The key to this result is that patents prevent independent creation.

A. One important differences between Patent and Copyright

There are some terminological issues that are best dealt with upfront. Copyrights provide ownership over creative works. Patents provide ownership over innovations. In principle, either could be set as long or as short as society would wish. In principle, either could allow exceptions to this ownership under certain conditions, restricting the owner’s ability to contract, as is the case, say, with compulsory licenses.

The economic argument for copyright and patent is based on promoting conditions that would allow creators or innovators to reap financial rewards for producing works and innovations. One
potential problem occurs when competition in the market prevents creators or innovators from covering the costs of invention or creation. Those who would copy an innovation or artistic work, (i.e., free riders not paying for the costs of creation), can produce their copies at lower average cost than the producers from whom they copy because these free-riders do not have the burden of incurring the costs of creation. This would tend to reward free riders at the expense of the creators or innovators which runs contrary to both the moral rights of the creators and the logic of wanting to induce creation and innovation.

Instead, the copyright and patent system is actually based on a welfare rule of thumb. The rule is that allowing private ownership and market transactions provides sufficient incentives to ensure that products with social value will be produced. That is one of the fundamental precepts behind reliance on markets and one of the lessons of history is that economic systems based on other rules of thumb have been far inferior. Like democracy, there are many possible imperfections in capitalism but it has performed far better at generating wealth than any of the alternatives, as many natural experiments have verified.

The moral rights of creators seems a much simpler proposition that strikes most people as fair: you should be able to reap what you sow. If revenues are generated from a book that I write, should I not be able to benefit from those revenues? Once these two views are accepted as reasonable, then the economic rule of thumb implies that we should allow creators ownership over their creations.

The dimension of ownership that is central to this essay has to do with breadth of ownership that is provided to the copyright or patent owner. In particular, copyright as it currently exists allows similar or identical creative works to compete in the market, as long as they are independently created and not copied. Copyright ownership is exceedingly narrow, although it is sometimes difficult to define the borders of the concept of “independent creation.” Patent law, as it currently exists, does not allow identical but independently created innovations. One of the great difficulties of real world patent rules is determining how broad the ownership rights should be over competing innovations.

This differential treatment of independent creation between in these two components of intellectual property laws is what causes errors of analysis when lumping both copyright and patent together.

Is this difference in the treatment of independent creation inherent in the nature of the two rights? This is difficult question to answer and I do not pretend to have the answer here.

For artistic works covered by copyright, the very narrow breadth of protection seems to be a natural element inherent in the nature of the works themselves. For simplicity, let’s keep to the world of books, the first copyrighted works. And we can ignore the length of copyright right now by assuming that it lasts forever. Finally, we can assume that copyright provides full ownership rights, so that activities that might be considered fair use are ruled out unless these activities acquire permission of the copyright owner. These assumptions, which I shall use later in the paper, can be considered under the rubric of “inherent copyright.” Under these circumstances, how wide could copyright in written words be, in principle, before the protection would become absurd?

If copyright were made for the first act of writing, meaning that the first author would then be able to control the entire market for written works, most everyone would agree that this was nonsensical. The first author would gain control over the entire market for written works even
though most later written works would have nothing to do with the original work. There would be no economic or moral linkage suggesting that the first author should receive payment for all later works written by others. Nor would it seem remotely possible that only a monopoly over the ownership of all later written works would be sufficient to provide incentives for the creation of the first written work. Remember as well that the first book written by the first author need not even be valuable on its own.

Next, we can imagine the first author in a particular genre of works, travel books, say, or fantasy, being given the copyright over the entire genre. All the same arguments just made can be made again. It is possible that some great work dominates a genre for many years, and that many authors are influenced by such great works, but those are not necessarily the first works in a genre. Authors may join scientists in standing on the shoulders of giants, but broad copyright ownership would likely be given based on time to market, not overall influence. If copyright ownership were based on the influence or quality of a work, instead of upon being first, there would need to be some method, no doubt controversial and somewhat arbitrary, to choose the most influential works and the nature of the control over the market given to these works. Although there would be some linkage for providing such market control based on moral rights and economic logic, the linkage would be weak. And the process would be very messy.

This messiness and arbitrariness is avoided by restricting the breadth of copyright protection to cover merely the work created (with derivatives, translations, and so forth providing a small amount of unavoidable messiness). The breadth of copyright is so narrow that it does not exclude other works that are virtually the same, as long as the new versions were created independently. In essence, copyright is giving the creator ownership over his version of the work. In other words, if an author put in all the work to create a second version of the work, without copying from anyone else (although there is some messiness in the meaning of “copying”), he is also given ownership over his version of the work and the rewards to the two creators are allowed to be whatever the market will provide. This is perfectly sensible from a moral rights perspective since there is no reason to prefer one independently created work to another. If the second work were to cut into the revenues of the first work, as would be expected, the answer to complaints about this competition would be “tough,” as my acquaintances in Brooklyn used to say. They would also say “if you don’t like it, create a work that cannot be duplicated.”

Of course, this works well in the real world because in reality neither a book, nor a chapter, nor a page is likely to ever be an independently created duplicate, because writing is a characteristic closely tied to an individual’s unique personality. Only short or small artistic works (such as a few dozen bars of music are ever likely to be independently duplicated. So we have a legal rule (independent creation being allowed) over an event that is unlikely to occur, therefore making it easy to adopt.

It is natural in copyright to allow independent creations to compete with the first work, because it is so unlikely to occur. Providing ownership through copyright is also an easy policy to recommend because it merely provides ownership. Because it allows independent creation it does not restrict competition in the creation of competing works. It mimics almost perfectly the template of how private competitive markets are supposed to work.

Patent law is different; patent law does not allow independent creation. It goes out of its way to prevent independent creators from being able to compete with the first creator to get the patent.
And unlike copyright, it is common for inventors to come up with the same or very similar technique for solving whatever problem their patent addresses.

Why does this key difference between patents and copyright exist? We can only speculate, but there are some good reasons to prevent independent but later innovators from competing with the first innovator (patent) whereas there are not good reasons to prevent independent but later creators from competing with the first creator (copyright). These reasons are purely economic in nature and not moral in any sense, since a later independent innovator would seem to have the same moral standing to an invention as an earlier innovator.

The provision of ownership over an innovation in order to allow the innovator to generate revenues and possibly profits seems straightforward enough, and as far as this goes, is similar to copyright. But if all independent innovators were allowed to bring their products to market, the profits generated by the group holding the patents would be less than the profits generated by a single patent holder if the patent holders competed with one another. It is possible that the profits (excluding costs of the innovation) generated by the shared patent owners might be high enough to make the investments worthwhile. Even if the profits were not sufficient, antitrust laws could be modified to allow the firms to cartelize and try to earn profits closer to monopoly level, although the profits would be shared among the independent creators instead of just the single first patent holder. Nevertheless, independent creation in patent, after the first, is not given any ownership rights.

Legal scholars may know why patent law disallows independent but later innovations, although I make no such claim to that knowledge. Later innovations may not be considered “novel” because the earlier version of the innovation removed the novelty, but this is a mere technicality.

A more serious problem with allowing independent creation in patent law is that it might not be possible to actually know whether the creation was actually independent. Clearly, such a determination would be difficult, but innovators could presumably hire firms specializing in verifying whether the later firm avoided copying the first innovator. This would be akin to reverse engineering a piece of software while trying to avoid copyright infringement by ensuring that their coders do not have access to the original code (i.e., a clean room design).

It is also possible that sharing whatever profits come from the patented item will leave each inventor with too little profit to make the innovative effort worthwhile. It might be thought that anticipation of such a division of profits would cause many worthwhile innovations to fail to be brought to market. Of course, the current system causes the second and later potential innovators to generate zero revenues, and uncertainty with regard to who gets the first patent on an innovation might also be thought to have equal or more deleterious effects on the incentive to innovate.

Regardless of why patents fail to allow independent inventors to share in the fruits of the innovation, I will assume from this point on that patents do not allow independent creation whereas copyright does. This will be the crucial difference that makes lumping copyright and patent in the same analysis questionable, particularly with regard to “monopoly” and “notice costs.”
B. Notice Costs

Notice costs have been defined as the costs that a property owner incurs in trying to demarcate the boundaries between his property and any adjoining properties. In the case of land, for example, fences can demarcate the boundaries fairly easily although at some cost since creating fences is not free. Notice costs can be incurred by those wanting to keep trespassers out (by building a fence, say) or those wanting to make sure that they do not trespass on someone else’s land (by looking for fences or other property demarcations).

It is claimed that notice costs are considerably higher for intangible properties such as copyrights and patents because they do not have the simple boundaries found in three dimensional physical properties. Here are two such claims:

- Compare land with intellectual property. Information costs are more significant in intellectual property than in real property and personal property law. Because they are intangible, determining and measuring the boundaries of intellectual goods are more difficult than determining and measuring the boundaries of real property. [Long, 2004, p 483]

- In contrast [to land], developers of many intangible resource projects face significant problems in identifying potentially conflicting rights, ascertaining the boundaries of those properties that they can find, locating the owners of potentially conflicting properties, and assessing the scope of potentially conflicting rights. [Menell and Meuer, 2013, p 18]

This logic seems correct because physical items need demarcation in only three physical dimensions, or in the case of land, two dimensions (the surface of the earth) whereas the “borders” of intangible properties are not limited by physical dimensions and so afford many more potential adjacencies.

One metaphorical dimension involved with intangible properties is that of time. Unlike land, which has a fixed location that does not change with time (although there can be important changes to quality or ownership over time), intangible properties have dates in which they come into existence and any potential problems with borders requires that the time periods between two properties match up.

There are two potential parties to any border or intersection of intangible creative or innovative properties.¹ There is the (intellectual) property that is created first and there is the intellectual property or use of an intellectual property that comes afterwards.

The owner of the early property might want to avoid having his property encroached by anyone who is the owner of a later property. In such a case, which we can think of as “rearview” notice costs, the costs are incurred by the original property owner trying to forestall encroachment by potential later property owners. Similarly, later owners might want to avoid encroaching any of the earlier property owners and thus may engage in receiving what we might call frontview “notice permissions,” which are attempts (negotiations, including payments) to make sure that no

---

¹ Long classifies “observers” dealing with the notice costs of intellectual products into three groups which do not include the creators of originals. My use of groups, although superficially similar to hers, is quite different.
encroachment of prior works has taken place without the permission of the earlier property owners. Each side wishes to avoid accidental encroachments because the costs of dealing with encroachments are pure deadweight losses borne in part by either or both of the parties (as well as society at large). Consumers of legally purchased intellectual products expect that all the necessary frontview permissions have been obtained by the seller of their copy.

Creators of intellectual products generally incur both rearview and frontview notice costs because at any point since the introduction of copyright and patent laws, there are already in existence past protected creations and there will undoubtedly be future creations.

In spite of its intuitive appeal, looking at the number of potential borders separating these properties may not be the best way to think about notice costs. With all the possible multiple intersections of various intellectual properties, does it really need to be the case that notice costs for intellectual properties are higher than for physical items, as sometimes claimed?

Although I cannot answer this question definitively, in the case of what I refer to as “intrinsic copyright,” defined in a footnote below,² it is quite easy to conclude that many notice costs are surprisingly low.

Why are the notice costs of intrinsic copyright so low? First, a simple rule reduces frontview and rearview notice costs to almost zero. The simple rule that society could adopt with respect to copyright is that any copying of a work is a violation of copyright, just as it is accepted that a student copying another student’s exam is considered cheating or that an empty automobile does not belong to just any stranger walking by. With this simple rule in place, copyright owners do not need to worry about whether later creators of copyrighted works are aware of the demarcations of their works. There are thus no rearview notice costs. Creators following this simple rule do not need to lookout for any demarcations. These costs are lower than, say, land boundaries. That is because, as I show momentarily, an infringer will always know that infringement is taking place, meaning that no creator can accidentally use part of an earlier copyrighted work in a manner that infringes the original work. In this situation the original owner need not spend any resources on demarcations otherwise intended to alert potential infringers about possibly infringing acts.

And why does the infringer always know when infringement is taking place? The answer, as the reader may have already divined, is that the potential infringer merely needs to determine whether he is creating a new work from whole cloth or whether he is copying the work or works of others, just as students know whether they are copying someone else’s answers on an exam. Because independent creation is entirely sufficient to create a legal property right under copyright, creators do not need to know the demarcation lines of any other prior works as long as they are not directly copying someone else’s work. I hazard to claim that almost everyone knows, at every moment in time, whether they are creating something from scratch or whether they are copying someone

² In an attempt to delineate the differences between copyright and patent, and to highlight how their disparate treatment of independent creation affects the analysis, I am use a term, “intrinsic copyright” or “intrinsic patent” to imply that copyright or patent has no limitation on its duration and no limitation on its ownership rights. Thus limitations such as fair-use or compulsory licenses are ruled out.
else’s work.³ An author of a work that is partly original and partly copied knows, at the time of creation, when copying is occurring and when it is not.

The only logical complication I can see would be a hypothetical person with a photographic memory who cannot distinguish between works sitting in memory and works being created de novo. There is also the possibility that someone might unconsciously know a music tune and then recreate it thinking that it is an original creation when in fact it was not. Obviously there could be other difficult examples as well. But in the big picture there must be few enough of these cases that we can essentially ignore them.

There are still notice costs for other categories of users of copyright works. An important backward notice cost that still exists in the case of inherent copyright would be the cost for those intentionally copying the works of others but wishing not to violate copyright. These individuals or organizations need to find a negotiating partner with whom to determine the price for permission to copy. These negotiation costs are where a registry can prove to be a beneficial, as long as the registry is updated when ownership over the work is changed.

These costs of finding a negotiating partner, the owner of the copyrighted good, are real. And they can be large in particular cases. I presume these costs can be considered to be notice costs. But it is not clear to me that these particular notice costs raise the overall notice costs for copyrighted items, on average, above that for tangible goods, such as land or automobiles. Particularly when most copyrighted works likely to have significant market value contain information about their creator or owner in readily noticeable locations, and most valuable works tend to be of relatively recent creation. It is true that there are orphaned works and other instances will occur where it is costly to find an owner, but the market value of these instances are not likely to be a large proportion of the overall market value of copying copyrighted goods.

Finally, there is one more class of copyright user that imposes costs on copyright owners in the case of inherent copyright. For example, some individuals may choose to copy someone else’s work without permission. This avoidance of payment will cause the owner of the original work to use resources to try to detect and/or deter unauthorized copying. These costs, however, are not notice costs, but instead are ordinary property enforcement costs. When someone intends to violate copyright, notice costs are irrelevant. There is no point in telling a thief that the item he is stealing is protected by a property right. He doesn’t care.

So let’s summarize the result for copyright. Except for the partnering costs by those intentionally copying another work, copyright does not impose notice costs on backward looking copyright owners because they do not need to worry about infringement caused by later authors uncertain about the boundaries of prior works. It does not impose notice costs on forward looking creators because all they need to do to avoid infringing is to avoid copying someone else’s work. Notice costs are only large for the case of intentional copiers who wish to follow the law. These costs are real, but so are the costs in determining the boundaries and owners of many tangible goods. Given that creators of many copyrighted items do not directly copy from other copyrighted items and that the costs of finding the entity capable of providing copyrighted permissions for valuable

³ This may be a bit of a simplification as cameras (or microphones) might inadvertently capture (copy) a copyrighted work of art in the background.
copyrighted products need not be very high, it is unclear that copyrighted goods have higher notice costs than most categories of tangible goods.

This result is contrary to the intuition mentioned above about notice costs being unusually high for intangible goods. I should also point out that the large number of copyrighted goods is also not relevant to a discussion of these notice costs, except for the costs to copiers of finding someone with whom to transact.

The reader will rightly object that copyright is not this simple in the real world, that “intrinsic” copyright does not exist in the real world, and I accept this point. I will return to the real world case in the next section.

But first I want to note that although copyright often has very low notice costs, patents do not have the same low notice costs. With respect to notice costs, copyright and patent are like night and day. This is the case even if patents share the two conditions required for “intrinsic” copyright: first, that the property right did not expire and second, that the property right was complete.

The reason that notice costs would still be high for intrinsic patents is simple. Backward looking notice costs exist for patents because patent owners want to reap licensing fees from later users of the technology, including those who might have been competing to be first to market. Going after inadvertent infringers is not as efficient as arranging negotiations before the innovation is used by others, and rational patents owners should be willing to undertake notice investments up to the point where the marginal gain from additional profits is equal to the extra costs.

Nor is there any expectation that other innovators will be able to avoid infringement at zero cost by following some simple production rule for innovation, the way that there was for creators under copyright who merely needed to create works de novo. Even innovating in something like a “clean room” environment does not protect the follow-on patent innovator from infringement. Thus, inadvertent infringements will occur and the backward looking innovator wants to reduce these infringements by putting up notices.

Forward looking notice costs are also positive in this case of “intrinsic” patents because innovators will wind up with no ownership over their innovation if they are not the first to patent the innovation. Thus they will want to ascertain, as best and as early as they can, whether previous innovators have already patented the innovation. They also will want to learn, as best they can, who else is working on similar innovations and how far along those competitors are. Discovering this information can help avoid the potentially very costly investment in an innovation that will not bring any return if someone else gets the patent. Thus expenditures will be made with the hope of identifying previous and current competitor innovations where the boundaries overlap.

---

4 This cost needs to be separated from the negotiated price of a copyright permission, which might be high but is not a notice cost.

5 Other writers (Menell and Meurer) have brought up the possibility of the patent holder hiding the patent and then holding up other innovators of substitute innovations after significant investments have been made. This is certainly possible, but as Menell and Meurer admit, such behavior depends on how damages are calculated and whether courts and legislators allow the system to be gamed in this manner. It is not necessarily a result that flows from the basic conception of patent law.
It is also the case that the exact demarcation of what a patent protects is less precisely defined than are the boundaries of a plot of land or other tangible good. Patent holders naturally want to claim the widest boundaries for their innovations and competing innovators want those boundaries to be narrower. Because there is imprecision in knowing where exactly the boundaries are, however, these costs are higher for patents than for many other properties, such as land or creative works covered by copyright. These are important costs. These costs, however, do not seem to be a form of notice cost, per se. They are costs imposed by the imprecision of the system, adjudication costs if you will. These are costs incurred before proper notice can even be given.

C. Notice Costs in the Real World

The “intrinsic” copyright regime described in the previous section is obviously not the actual copyright system that exists, although it seems a feasible system if governments were so inclined to set one up (other than the U.S., where the constitution limits the duration). Actual copyright laws have limited ownership rights in terms of duration, and limited rights in terms of the breadth of the ownership. Fair use, for example, was a defense against copyright infringement in the common law that was codified into statutory law. Although the four factors listed as being important for determining fair use provide some guidance, it can be very difficult to know when an act of copying will be considered fair use.

Do these real world restrictions on copyright’s duration and breadth alter notice costs?

These restrictions on copyright’s duration and breadth certainly induce higher transactions costs on users of copyright works even though users, it should be remembered, presumably are made better off by these restrictions on copyright since these restrictions reduce the number of potential infringements, eliminating the need to make payments for the copying of copyrighted goods. Although individuals creating works de novo still have no forward looking notice costs, individuals copying works or parts of works now face uncertainty in knowing whether permissions are legally required. These copiers, instead of knowing with certainty that they are violating copyright if they do not acquire permission, as they did in the hypothetical intrinsic copyright of the previous section, might now not be violating copyright if the copying is considered fair use or if copyright has expired.

The limitation on copyright’s duration, for example, imposes costs on users of the work when they expend resources to determine whether copyright is still in force. The copyright owner will wish to provide notice of ownership during the life of the copyright, but has no incentive to provide notice of when copyright has expired.

When copyright has expired, the work enters the public domain. Since no one can legally be kept from using a public domain work, there is no one to “fence off” the boundaries. When a work is unowned, who is it that even could provide “notice”?6 Third parties, possibly the government

---

6 If the public domain were owned, we would expect its owner to provide notice of the works contained within, just as we would expect a public park to provide information that it is a public park. If some organization, even the government, were to create a registry of all works in the public domain, it could charge a small fee to all users of the material contained within and indemnify anyone making use of the works against charges of copyright infringement. Similarly, if all works under copyright were kept in an updated register (or if copyright were a fixed duration), it would presumably be of low cost to determine whether a work was still under copyright.
could provide such information, but their motivation would be different than an owner’s motivation.

The breadth restrictions on copyright are even less conducive to a notice-centric view of the world. If a work is still under copyright, any intentional copier trying to follow the law would need to acquire the rights or incur the costs of determining whether the copying constitutes fair use or not. The costs of determining whether an act of copying is definitively fair use seem likely to be quite high.

Is the cost of determining what constitutes fair-use a notice cost? Menell and Meurer (2013) certainly classify it as such. Taking but one of several examples:

Somewhat distinct notice problems arise in the development of expressive works. While tangible resource boundaries tend to be well-defined and capable of precise measurement, the scope of copyrighted works and the permissible extent of fair use can be difficult to ascertain. [my emphasis]

I am in agreement that fair use imposes costs and uncertainty on the users and producers of copyrighted works. But I question whether the copyright breadth (fair use) issue is a notice problem. There is no notice that the copyright owner can give that will mitigate the costs to the copier of determining whether or not an act of copying is fair use for the simple reason that the copyright owner does not know the answer. There is no notice that the copier can give to the copyright owner that would lower the owner’s cost of determining whether the unauthorized copying of his work is illegal or not. If there is no form of notice that can lower these costs, it seems to me to be unreasonable to refer to these costs as notice costs or notice problems. Notice costs would then be just another name for transactions costs or information costs.

Let us also be clear about the actions that led to this increase in information costs relative to inherent copyright. Society imposed these costs by restricting the ownership rights of copyright owners. There is an implicit expectation that these extra information costs found in copyright laws allowing fair use and shorter durations are more than compensated for by the benefits to society from allowing copiers to legally avoid purchasing the rights to a copyrighted work under certain conditions. When the decision to allow fair-use was made, these extra uncertainty costs should have entered into the overall analysis that led to the decision.

**D. The “Monopoly” in Patent and Copyright**

There is another fundamental difference between copyright and patents that is often ignored, if not misunderstood. That has to do with the nature of the “monopoly” provided by copyright or patent.

I have often noticed hostility to the term intellectual property but had always considered that hostility to be against the provision of a right of property to these creations. However, the unwillingness of many analysts to understand the nature of the “monopoly” associated with copyright makes me unsure of this interpretation. It is possible that these critics of intellectual property actually misunderstand the distinction between monopoly and property.

A property right does provide a nominal monopoly over the item covered by the property right. I have monopoly control of the automobile that I own, over the smartphone that I own, and over my
person, which can be thought of as self-ownership. Ownership allows the owner and only the owner (subject to legal limitations) to determine how the item in question is used.

But does anyone believe that this nominal monopoly, due merely to ownership, provides an actual economic monopoly? Does my auto provide me an economic monopoly in transportation, or my cell phone provide me an economic monopoly in telecommunications? Of course not. Nor does the “ownership” over their personal talents necessarily provide most workers with monopoly power in the labor market.7

That is not to say that ownership cannot help to bring about the exercise of monopoly power. If I own the property rights to most of the world’s diamonds, then my ownership rights in diamonds allows the exercise of monopoly power in the diamond market. Legal ownership, or some other similar form of control, is a necessary condition for the sale of products. A monopoly sells products without the hindrance of competitors. Being able to sell the diamonds is a requirement for this monopoly, but it is not the key ingredient leading to monopoly. Having all or almost all the diamonds is the key ingredient.

The key requirement for the economic exercise of monopoly power for product A is that there are few or no competing items that consumers consider to be good substitutes for product A in a particular market. The existence of substitutes in a market implies that competition will drive prices below the level that would exist without the presence of competing substitute products and that monopoly power will be diminished or eliminated.

How does this understanding apply to copyrights and patents?

It seems clear that copyrights and patents both give ownership of an intangible work or innovation to the party receiving the right. But copyright provides only ownership. Patent, by way of contrast, provides more than ownership because independent creation is prohibited. Patents provide a limitation on the technologies that can be used to compete with the patented technology, limiting competition for the products produced with the innovation. Patents can create a monopoly when the breadth of a patent restricts most other related technologies.

Obviously, the breadth of patent protection depends on the behavior of the patent office and courts. Some patents will confer monopoly power and others will not. A patent is not a sufficient condition for the provision of monopoly power, but it has characteristics that enhance the possibility of providing monopoly power. That is because even when there are many other similar, independently created innovations with the potential of competing with the patented products, those potentially competing innovations are likely to be preempted by the patent. As already noted, copyright does not preempt competing but independently created works.

This distinction between simple ownership versus ownership that also restricts competition is what separates copyright from patents. The different treatment of independent creation between copyright and patent is what causes the difference in whether a monopoly is brought about by the granting of patent or copyright, just as the different treatment of independent creation was at the heart of the difference in notice costs between the two. The fact that patent law’s broader scope

7 I discuss in more detail the distinction between ownership and monopoly in my 2015 George Mason Law Review article.
enhances monopoly relative to copyright does not imply that patent law should be changed to remove this monopoly enhancing element. A full discussion of that issue is beyond the scope of this paper.

The distinction between ownership and monopoly in the case of copyright is one that has been made many times, over many generations, by many individuals. A relatively early mention comes from Thomas Huxley (“Darwin’s bulldog”), in 1877:\(^8\)

\[
\text{[S]o that in my apprehension the application of the word "monopoly" to persons who possess rights under the copyright law is an entire mistake; it is merely a contrivance arising out of the peculiar nature of book property, to put that property upon the same footing as other kinds of property.}
\]

Edmund Kitch complained in 2000 that economists and lawyers were still failing to make the distinction between ownership and monopoly in the case of copyright, what he called “elementary and persistent errors.” This distinction continues to be lost upon many, such as Boldrin and Levine in their 2008 book, so aptly titled “Against Intellectual Monopoly” in which the term “monopoly” is misapplied and copyright and patent law are lumped together as one.

So why do academics writing about copyright frequently repeat the claim that copyright is a form of monopoly? Is it a real belief, or is it a contrivance? It is quite possible to be against the assignment of a property right for creative works without resorting to false claims of monopoly. But it seems likely that the rhetoric of the case against copyright appears stronger if the term monopoly can be bandied about.

Assuming that those copyright critics who claim that copyright creates a monopoly are not just looking to score debating points, then it seems that they must be looking not at the nature of the rights created but instead at the market where copies of works are sold. Looked at in this way, particularly through the lens of models found in economics textbooks, the market for a creative work might superficially look like a monopoly.

That is because the demand curve for individual titles is downward sloping. The seller of a particular novel, for example, is a price searcher, not a price taker. Any product with a downward sloping demand looks superficially like a textbook monopoly. Any profit maximizing firm facing a downward sloping demand curve will produce an output where price is not equal to marginal cost and will not be on the bottom of the firm’s average cost curve.

This leaves us with two questions. Are all firms facing downward sloping demands deserving of the term monopolist? Is there a welfare gain to be had by increasing output for those units where the demand curve is above the marginal cost curve? After all, we teach that these missing units (the unproduced units where the willingness to pay is greater than the marginal cost) are responsible for deadweight losses in the case of monopoly. So doesn’t that hold here?

The answer to both questions is that it depends. Let’s examine another simple textbook case. Let’s take a purely monopolistic producer, the only firm in an industry. Let's assume that the profit

---

\(^8\) Found in paragraph 5553 in the testimony to the Royal Commission on Copyright.
maximizing output just allows the monopoly firm to just break even. The textbooks will tell us that this is a product that can only exist in the case of a monopoly. Figure 1 represents that case.

If this industry were monopolized, on the other hand, producing at the profit maximizing output Q, the monopolist could survive, just covering all costs, including opportunity cost, and thus earning a normal return on investment. Since consumers earn a surplus on the Q units that are produced by the monopolists, outside observers would applaud this monopolist for producing this product and benefitting society. This situation is known as an instance of natural monopoly.

The welfare characteristics of this monopoly solution are somewhat difficult to classify. The output from Q to Q* has a marginal cost of production which is below the demand for those units. Society would be better off if we could get those additional units produced. But at this point we encounter Demsetz’ (1969) nirvana fallacy. If the only actual choices are between competition and monopoly, then Q is the efficient solution. It doesn’t matter whether we could imagine a better

\[9\] This is a horizontal summation of marginal cost curves, and the only portion of the curves which are summed are those portions which lie above the related average variable cost curve.
solution if that better solution is not feasible. Assigning the government the task of achieving the better solution, if it is not realistic, is the nirvana fallacy. The ideal solution, if unachievable, should not be classified as the efficient solution.

Let’s segue back to the book market in Figure 1. If the book publisher had the super-human capabilities of being a perfectly discriminating monopolist, then a larger number of copies of the novel \( Q^* \) would be produced than would be the case for the ordinary monopolistic publisher, and overall economic welfare would be increased. But a perfectly discriminating monopolist is a textbook fiction, not a feasible solution, so \( Q^* \) cannot be achieved that way.

Most firms producing most products in most industries face downward sloping demand curves. That is because most firms in an industry produce products that are not identical to one another. Do we accuse most firms in most industries of being monopolists? We do not. What is going on in the many industries that are competitive but where products are differentiated from one another, meaning that different products in the same industry are imperfect substitutes?

Economists have developed an archetypical model to account for these circumstances—monopolistic competition. The key conditions of this model are that competition drives the profits to zero for all producers of differentiated products in a market and that the differentiation means that each producer faces a downward sloping demand curve. The diagrammatic representation of this model for the typical firm, as found in textbooks, is identical to Figure 1.

The textbooks will also tell you that result in this model (monopolistic competition) is not ideal because output \( Q^* \) would be better than output \( Q \), and because the firm is not at the bottom of its average cost curve (AC). But as we have seen, not being ideal is not the same as being inefficient. And there were very strong debates within the economics profession about whether \( Q^* \) really was better than \( Q \) since \( Q \) was associated with zero product differentiation, and differentiation, otherwise known as variety, has its own value which is missing from an analysis of figure 1. This deviation from the perfectly competitive ideal led to criticisms of advertising, of fashions, of model year changes in products, of variety. Over time, a more complete understanding of the tradeoffs involved with variety took hold and the more doctrinaire concerns with the so-called inefficiency in Figure 1 that the apparatchiks of the Soviet Union and many economists in the 1930s and 1940s believed, was discarded. We are still allowed to wear clothing of different styles and colors.

Applying this model to markets for creative works requires separating the creative works into two distinct elements. On the one hand, you have the market for titles, say all fictional mystery books. They are substitutes for one another but also differentiated from one another. The market for titles would be expected to have a downward sloping demand\(^{10}\) and an upward sloping supply.\(^{11}\)

---

\(^{10}\) The construction of the demand for titles is quite different than the standard demand curve found in economic textbooks for normal markets. That is because creative works are instances of non-rivalrous goods (sometimes unfortunately referred to as “public goods”) and such goods have the demands on the part of consumers summed vertically rather than the normal horizontal addition found in textbooks for rivalrous goods. The vertical summation is also tricky because the number of consumers whose demands need to be added together depends on the ability of the sellers of titles to price discriminate.

\(^{11}\) Some analysts have claimed that the supply of creative works is not upward sloping in the sense of requiring higher payments to produce a greater number of works. This is an empirical issue, but although I agree that some works will be created without any compensation to the creator, I seriously doubt that the number of works will not react
Titles have to be different from one another. What kind of book market would it be where all the books were the same and where consumers would read the same book over and over again, the way they might drink the same coffee over and over again?\textsuperscript{12}

The producer of any single book title also faces a downward sloping demand for the copies of the book title being sold. All units of a single title are essentially perfect substitutes for one another and fit the standard basic model of markets. Figure 1 represents the market for an individual title.

In Figure 1 the producer of the typical title is shown earning zero economic profit. The vast majority of authors and other creators produce works that tend to have many close substitutes. Unlike the citizens of Lake Wobegon, the talent of most creators is not sufficiently above average for them to experience monopoly power in the markets in which these works are sold. Competition in the market for titles will tend to drive profits down toward zero for the typical book. The ownership that copyright provides does not allow these creators to wield any monopoly power.

Unlike the model, some varieties are better than others. Copyright allows individuals with unique (monopoly) skills to exercise their monopoly in markets. Stephen King and Bruce Springsteen have monopoly talent in the markets for books and music respectively. Copyright provides Mr. King and Mr. Springsteen with the ability to access their innate monopoly power, and to exercise that monopoly power in the market for horror books and music respectively. They earn monopoly rents in the same way that basketball’s LeBron James and golf’s Jordan Spieth do.

Copyright provides the ownership over their works that allows their monopolistic talent to generate rents. It is still incorrect to assert that copyright is the source of the monopoly power for these creators. For this small class of titles (which would have an outsized share of the total sales market), it is possible to argue that feasible welfare gains could come about by weakening copyright (ownership). In these instances, allowing free riders to share some ownership of the works (and print competing copies of the work) to move the market toward \( Q^* \) could increase welfare. Of course, the same could be said for every firm earning above normal rents.

But the source of the monopoly is not the copyright. These inefficiencies due to innate talent are not the target of antitrust activities elsewhere in the economy, and it is unclear why they should be disallowed for creators when they are allowed in virtually all other labor markets.\textsuperscript{13}

\begin{footnotesize}
\begin{enumerate}
\item[12] Although it would be possible in principle to have book readers always rereading the same title over and over again, the way that children can listen the same story or watch the same movie over and over again, few rational adults would claim that such a market would be welfare maximizing. Variety obviously has value. Yet the ideal state of perfect competition, where producers are at the bottom of the average cost curve and where price equals marginal cost, can only occur in this stultified world (or in the equally idealized world of perfect price discrimination—the word “perfect” is the tipoff).
\item[13] On this point see my paper in the George Washington Law Review. It is the case that if free-riders were allowed to impinge on the ownership of one of the titles written by an author with inherent monopoly power, the free riders could lower the price of the book, reduce the profits generated by the monopoly talent, and increase consumer welfare. But just because this can be more easily undertaken in this market, in most other markets monopoly power is allowed to be exercised with impunity and it is unclear why copyright owners should be singled out.
\end{enumerate}
\end{footnotesize}
The fact that book titles are not identical, and that they have downward sloping demands may be the reason that some copyright critics claim that copyright provides a monopoly. But as we have seen, applying the term “monopoly” in those circumstances is just an instance of sloppy thinking.

E. Conclusions


The Royal Commission on Copyright, Minutes of the Evidence Taken Before The Royal Commission On Copyright, (1878).

Plant

Waldfogel