

SUBSTITUTION AND SCHUMPETERIAN EFFECTS IN THE LIFE CYCLE OF COPYRIGHTED WORKS

Ariel Katz

Assistant Professor
Centre for Innovation Law and Policy
Faculty of Law, University of Toronto
ariel.katz@utoronto.ca

To be presented at the
6th Annual Intellectual Property Scholars Conference on August 10-11, 2006.

Preliminary draft

Comments are welcome

Please do not cite without permission

I. INTRODUCTION

Copyright law tries to manage several trade-offs. In addition to the well known tension between incentive and access that follows from the grant of exclusive rights in creative works, copyright law seeks to reconcile the tension between the encouragement of creative pursuits by granting exclusive rights to authors, and the concern that doing so would limit the ability of follow-on creators “to produce new works by building on the ideas and information contained in the works of others.”¹ Several copyright doctrines, rules and exceptions, such as the originality requirement, the idea-expression dichotomy and fair use are said to address this tension. At one extreme, a prohibition on making *identical* copies of a work maintains the incentive to invest in new works while imposing no cost on future creators, as by definition, such a copyist creates a perfect substitute for the work, which directly competes with the original author, without adding anything

¹ CCH Canadian Ltd v. The Law Society of Upper Canada [2004] 1 S.C.R. 339.

creative of her own. But as we move further along from outlawing not only identical copies to outlawing copying of a ‘substantial part’ of a work, the cost of suppressing future creativity increases. But also, as we move further away from the identical copy, the new work, containing its own creative content, becomes more differentiated from the first. Possibly, if the new work then negatively affects the demand for the old one, it is because it becomes superior to it. It affects the profits derived from the first not (only) because it grabs market share and reduces the first work’s price, but because it provides something new that consumers value. Its capability to supplant the old work represents a form of Schumpeterian competition, in which superior innovative products displace the older ones because of their superior value. Outlawing the new work in such cases would place a heavy toll on follow-on creativity and on the social welfare resulting from it.

So generally, while *perfect* copies’ only effect is to reduce the first works’ profitability through substitution and without creating any additional value, we can observe two effects created by *imperfect* copies: works that copy parts of older works but modify them to create something else. Imperfect copies may harm an existing work through a substitution effect and through a Schumpeterian effect. Both negatively impact the first work but for different reasons. Copyright law is (or should be) interested in addressing pure substitution effects without discouraging Schumpeterian effects.

While it is probable that in many cases both effects co-exist, I am particularly interested in exploring the possibility that the importance of each effect changes along the life cycle of a creative work. My purpose here is not to draw the proper line between unlawful “imitation” and permissible “improvement”,² or between “transformative” and “superseding” uses,³ or whatever other terminology may be used, but rather to highlight the effect of time on where this line should be. A copyrighted work is born when an idea is conceived and initially expressed and fixed; it is then brought to the market and matures. Sooner or later its glory days elapse (perhaps with some chances for potential comeback), until it is ultimately forgotten in the archives of cultural relics. At each period, the relative importance of substitution and Schumpeterian effects created by

² Mark A. Lemley, *The Economics of Improvement in Intellectual Property Law*, 75 Tex. L. Rev. 989 (1997).

³ *E.g.*, *Campbell v. Acuff-Rose Music, Inc.* 510 U.S. 569, 579 (1994).

copies may vary. For example, at the early stages of a work's life, its value (or at that stage the potential value) may lie predominantly in the ideas expressed therein. At this stage, other works embodying the same ideas but expressing them in a slightly different manner can function as very close substitutes. The commercial viability of the work at that stage is highly susceptible to competition from other similar (even if not identical) works. As the work matures, however, its commercial success tends to depend less on the intrinsic value of the ideas and their specific expression and more on creative inputs of other co-producers and endogenous factors, such as the word of critics and the work's ability to represent shared identities of groups' members or become a shared cultural focal point for meaningful social interactions. At that stage, the work is much less likely to be susceptible to substitution effect, although its popularity may make it attractive for follow-on creators to build upon, thus increasing its exposure to Schumpeterian effects.

At both stages the copyright owner of the first work might invoke her copyright to fend off the competitor, or have a claim in her profits, but while at the early stage fending off the competitor or recognizing the claim against her may be beneficial because it preserves the incentive to invest, using copyright law to protect against Schumpeterian effects may be socially undesirable.

If the relative importance of each of the effects changes over the life cycle of the work, then it may be worthwhile asking whether and how the law should take account of this dynamic. In two recent articles Professors Justin Hughes⁴ and Joseph Liu⁵ (separately) suggested that time should be a factor taken into fair use analysis. In this paper I will build on their insights and expand them, using the distinction between substitution and Schumpeterian effects to explore its applicability in other copyright areas. I will show how these insights can explain some of the details of the existing mechanical compulsory license for sound recordings, how they can inform the debate on the ability of copyright owners to use contract and technology to change the original allocation of entitlements as set by the copyright act and contract around users' liberties and the public domain, and how they can illuminate the idea-expression dichotomy.

⁴ Justin Hughes, *Fair Use Across Time*, 50 UCLA L. R. 775 (2003).

⁵ Joseph P. Liu, *Copyright and Time: A Proposal*, 101 Mich. L. Rev. 409 (2002).

The article proceeds as follows: Part I sets the distinction between substitution and Schumpeterian effects; Part II describes the life cycle of creative works and demonstrates how at different stages the relative magnitudes of each effect vary; Parts III, IV, and V apply this paper's insights may to the mechanical compulsory license, the debate on contracting around users' liberties and the public domain, and idea/expression dichotomy, Part VI offers an improved version of Hughes' and Liu's analysis of time and fair use; Part VII concludes.

II. SUBSTITUTION VS. SCHUMPETERIAN EFFECTS

Incentive theory of copyright easily justifies the prohibition on copying: without copyright protection, others' ability to copy and distribute copies would cause prices to drop quickly to the marginal cost of production and distribution and authors, unable to recoup the investment they incurred in creating the work would refrain from investing in the first place.⁶ By giving the author an exclusive right over the making and initial distribution of copies, the author can determine the number of copies made and set their price above marginal cost and generate enough profit *ex post* to cover the initial investment and account for the risk of failure to make render the investment worthwhile *ex ante*. The assumption here is that consumers (or a large enough number of them) view the unauthorized copy as a perfect substitute to the authorized copy, or at least good enough substitute. The existence of unauthorized copies harms the copyright owner in two related ways: first, because he no longer determines the number of copies available, the quantity of available copies increases, their price decreases and so do profits. Second, whatever profit the work does generate, they accrue not only to the owner but also to the copier.⁷

⁶ See e.g., William M. Landes & Richard A. Posner, *An Economic Analysis of Copyright Law*, 18 J. Legal Stud. 325 (1989), 333

⁷ Economic theory of copyright provides a more nuanced story however. First, it identifies that unauthorized copies do not always harm the copyright owner, see e.g. ... Moreover, even when unauthorized copying harms the copyright owner it does not necessarily follow that the incentives to create in the first place are equally harmed, see e.g. ...

As I have noted elsewhere,⁸ from an investment perspective, an author must also consider the possibility of competitive entry by less direct copiers or imitators. Copyright law provides partial solutions to this concern by providing that two works may not be literally identical and yet, for purposes of copyright infringement, be found to be substantially similar.⁹ But this provides only a partial solace for the author. Theoretically, competitors may create works that from consumers' perspective are very close substitutes. They may create with impunity identical works if they do so independently (i.e., without copying), and they may create functionally equivalent works differentiated enough to fall beyond the first work's protected area. If that happens, the price of both works may drop towards marginal cost exactly as in the case of unauthorized copies. Moreover, a prospective author contemplating investment in a new work must anticipate an even worse scenario, the *post* entry introduction of a non-infringing work that is sufficiently superior to her own that will totally displace it, not only compete with it.¹⁰

Copyright law, of course, affords no protection against such risks,¹¹ and consequently rewards and encourages the creation of works that add enough unique value so that they can displace previous ones, and stay long enough on the market until they ultimately are displaced by others, often long before the copyright legally expires.¹² In this sense, by prohibiting competition from identical or "substantially similar" copies, but allowing total destruction by other creative works, copyright law encourages the kind of competition described by Joseph Schumpeter as the "perennial gale of creative

⁸ Ariel Katz, *Making Sense of Nonsense*, at ____.

⁹ Melville B. Nimmer & David Nimmer, *Nimmer on Copyright*, Release 63 §13.03[A] (2004).

¹⁰ Katz, *supra* note 8.

¹¹ Stephen Breyer, *The Uneasy Case for Copyright - A Study of Copyright in Books, Photocopies, and Computer Programs*, 84 Harv. L. Rev. 281 (1970), at 327 ("copyright law is intended only to protect an author from competing sales of his own works; it was not designed to limit competition from the works of others.").

¹² Unique value of course, is not a prerequisite to copyrightability. The law only requires that a work be "original", which means that the work is not copied and, depending on the jurisdiction, that it represents a "modicum of creativity" (*Feist*, US), "skill and judgment" (*CCH*, Canada). In any event, despite differences in nuances, the amount of creativity required in almost all jurisdictions is rather low, *see* Gervais, *Feist Goes Global...* But although all original works are de jure equal, the incentive to create works providing unique value comes from the fact that only those which provide unique value will be rewarded by the market.

destruction”,¹³ in which new products displace old ones, and which themselves will be later displaced by the next generation of products. In this process the threat to firms comes not from close substitutes but rather from “the new commodity, the new technology, the new source of supply, the new type of organization ... competition which commands a decisive cost or quality advantage which strikes not at the margins of the profits and the outputs of the existing firms but at their foundations and their very lives.”¹⁴

While Schumpeter believed that “creative destruction” is “the essential fact about capitalism”¹⁵ generally (in fact, his analysis barely mentioned intellectual property)¹⁶ markets for copyrighted works usually exhibit this tendency of new products to rapidly displace old ones. With a few exceptions, creative destruction typically happens long before copyright expires and substitutive competition from identical copies becomes legal.

Note that in both substitutive and Schumpeterian competition the incumbent firm (for our purposes the owner of copyright over the existing work) may experience decline in sales and profit and may be forced to respond by lowering prices. The effect on the market for the incumbent’s work therefore does not indicate which type of competition he is facing. But while in the case of substitutive competition prices charged by both go down (resulting from an increase in supply of the same product),¹⁷ in the case of Schumpeterian competition, because the entrant provides a different, perhaps superior product, she may charge a higher price which previously only the incumbent could. Thus the prices charged by the entrant, rather than the harm to the incumbent can give us a better clue as to which type of competition the new work creates.

¹³ Joseph A. Schumpeter, *Capitalism, Socialism and Democracy* 84 (3rd ed. 1950).

¹⁴ Schumpeter, *id.*

¹⁵ SCHUMPETER, *id.* at 83.

¹⁶ Mark Blaug, *Why Did Schumpeter Neglect Intellectual Property Rights?*, 2 REV. ECON. RES. COPYRIGHT ISS. 69, 70 (2005).

¹⁷ It doesn’t mean that the incumbent necessarily lower his prices. Instead, he may decide to keep prices at their pre-entry level and sell only to the subset of consumer who have stronger preference to his product (i.e., consumers who are brand-loyal). This happens often in pharmaceutical markets, when upon expiry of a patent and entry of generic firms, brand name drug companies do not match the prices of generics but instead give up market share and focus on brand loyal consumers, *see*, William M. Landes & Richard A. Posner, *The Economic Structure of Intellectual Property Law*, (2003), Ch. 11.

The distinction between substitutive and Schumpeterian competition may be conceptually clear, but in practice both effects can happen simultaneously, as consumers may have varying tastes and preferences. Some consumers may be relatively indifferent as to which work they actually purchase, provided that it is within some genre, whereas others may have strong preferences to either the existing or the new work. But as a general proposition, substitutive competition causes prices of both works to go down, whereas Schumpeterian competition affects primarily the incumbent's prices. My purpose, though, is not to provide here the tools to distinguish between the two, but rather to advance the proposition that the even if both effects can coexist at any given time, their relative importance varies across time. I will submit that at earlier stages of creative works' life cycle substitutive effects dominate, whereas Schumpeterian effects dominate competition from non-literal copies of mature works. Because I assume that a copyright system guided by the public interest would be interested in discouraging substitutive competition but not Schumpeterian competition, recognizing that these effects vary over time matters. It will now be useful to discuss the life cycle of creative works and how they are affected by the different types of competition across time.

III. CREATIVE WORKS' LIFE CYCLE

Because this paper is interested in how different types of competition affect creative works across time, it may be useful to distinguish between a copyrighted "work" in the meaning of copyright law, and a copyrighted "product", as market forces affect only the latter. The two do not necessarily overlap. A copyrighted work is born when an idea is conceived and initially expressed and fixed in one of the forms recognized by the act (literary, artistic, musical, etc.). At this stage we have a "work" protected by the copyright act, but we don't necessarily have a "product". Perhaps more precisely, the product that we have at this stage will evolve: a sketch may become a picture presented in a gallery, a draft may become a published book; a play may become a show, a script a movie, and a musical composition a performed, recorded and released sound recording. What typically transforms the initial copyrighted "work" into a "product" finally put on the market is a process of co-production in which many additional elements are added to

the work, improve and develop it. Some elements may be added by the same author, and other by co-creators; some are creative while others are less. These may include additional copyrighted elements (e.g., music added to lyrics, to which an orchestral arrangement and sound recording are added), or not-copyrightable inputs such as the opinion of reviewers or editors, instructions given by a director, the charisma of a performer, some business decisions of a producer.

But the evolutionary process which ultimately determines the commercial success of the copyrighted product doesn't end at this point. After the product is put on the market its value may continue to grow as a result of inputs created by others. Again, some are more creative than other; some may be copyrighted works of themselves, while others aren't. Many markets for creative works exhibit "network effects" in which the value of the work depends not only on its intrinsic qualities, but also on the number of other people consuming it. Well documented in the case of information technology and software products,¹⁸ similar effects exist in the case of cultural goods whose value too increases with their consumption by other people. This happens because part of what people derive from cultural goods are relevant social interactions;¹⁹ "we do not want to read books nobody else reads, we do not want to see movies nobody else sees. We want to discuss, rave, slaughter and define ourselves by the things we like."²⁰ As in the case of information technology, cultural goods may have direct network effects, whereby the value of the work increases with the number of other people consuming it because of the opportunities to interact and exchange views with other people (just as having the same software or hardware platform allows consumers to exchange files with each other), and indirect effects, whereby the increased number of users induces the development of complementary products and services (e.g., compatible software and hardware, support services and literature in the case of information technology; commentary, critique, parody, academic courses, fan clubs, in the case of cultural products). As consumers'

¹⁸ See e.g., Ariel Katz, *A Network Effects Perspective on Software Piracy*, 55 U. Toronto L. J. 155 (2005); Mark A. Lemley & David McGowan, *Legal Implications of Network Economic Effects*, 86 Cal. L. R. 479 (1998).

¹⁹ Cass R. Sunstein & Edna Ullmann-Margalit, *Solidarity Goods*, 9 J. POL. PHIL. 129, 138 (2001).

²⁰ Martin Kretschmer et al., *Increasing Returns and Social Contagion in Cultural Industries*, 10 BRIT. J. MGMT. 61, S63 (1999).

decision which of two otherwise similar products to choose depends on how many other people have already chosen them, being the first to come up with a product and having an installed base of consumers as early as possible may be determinative for the products success.²¹

Another important attribute of creative works is that their quality is often hard to discern before consumption, and sometimes even after, therefore consumers' choices will depend in part on what other people think, and how many of them think so²² and on the reputation of the author for having provided works of certain quality in the past. "The main reason that we read the *Wall Street Journal* today is that we've found it useful in the past."²³

Sooner or later, however, the glory days of even the most successful works elapse, and with the exception of few works that become "classics", most works would be ultimately forgotten in the archives of cultural relics. This may happen way before the work formally falls into the public domain, as a result of the process of "creative destruction" described above.

This life cycle implies that at different points across it creative works would be vulnerable to different types of competitive threats. To identify these threats, let's identify a few relevant time points. Let T_1 be the moment of when an idea is initially conceived, expressed and fixed to constitute a copyrightable "work"; T_2 when the work-turned-product is released; T_3 when the product matures and reaches the peak of its commercial success; T_4 when it becomes obsolete but copyrighted; and T_5 when copyright expires the work falls into the public domain. While literal copying might expose the work to substitutive effects at every period,²⁴ non-literal copying might have different effects at different stages, at some points it may threaten the work by plain substitution, whereas in others it may threaten the work by "creative destruction".

²¹ Network effects do not always unambiguously increase the value of the product. Some network may experience congestion as the number of users increases beyond a certain point, *see* _____; other products may lose value beyond a certain level of use as part of their appeal is in their uniqueness, *see* _____.

²² *Id.*

²³ Carl Shapiro & Hal R. Varian, *Information Rules: A Strategic Guide to the Network Economy*, (1999), 5.

²⁴ Unless such forms of sharing-copying-piracy actually increase the demand, as mentioned above, *supra* note 7

T₁ – The Prototype

At T_1 the work competes against many other works in the market for publishing; it competes for an opportunity to enter a stage of transformation from “work” or a prototype of a product into a complete “product”. Two features make this market highly competitive. At T_1 the work’s value seems to lie predominantly in its intrinsic qualities and often in the ideas contained therein. But mere ideas aren’t protected and the intrinsic quality of the work is often hard to discern at this stage, either because most people don’t possess the necessary expertise, or because ultimately, the quality would depend on what other people think about the work. Therefore, at T_1 the work has to compete vigorously against many other works in the genre: not only works that may express similar ideas but also works expressing different ideas but serve a similar function. An example of the high substitutability of works at T_1 with which readers of this paper may be familiar is the typical statement in e-mails from law review editors rejecting the submission, apologizing that they must reject many fine articles from the hundreds submitted. As a result of this highly competitive nature of the market at T_1 , non-literal copying of the work can be as devastating as flagrant literal copying. In the eyes of publishers who consider which work to publish, there isn’t much difference between the authentic work, a literal copy of it, or a non-literal copy thereof. As prototypes of products yet to be developed, they all function as pretty close substitutes.²⁵ This means that substitutive effects dominate T_1 , and there is less concern that copyright protection would stifle Schumpeterian competition.

T₂ – From Prototype to Product

At T_2 the competitive dynamics change. Additional elements and creative inputs transform the work into a product with higher potential value. The value of the product now stems not only from the ideas the work conveys but also from the combination of the work’s intrinsic value and the value added by co-producers. These increase the uniqueness of the work and decrease its substitutability relative to other works against

²⁵ The name and reputation of the author, of course, may create a big difference in allowing the works of some authors to stand out among the many substitutes. This may imply that at T_1 an attribution right or trademark-like rules may suffice. However, these may be more effective in protecting known authors than unknown ones. Copyright therefore may help entry by non-famous authors.

which it competed at T_1 . Yet the actual quality of the work is still unknown and similar works within the same genre may still substitute it quite easily.

To clarify the difference between T_1 and T_2 consider a pop song. At T_1 the work is the musical composition: lyrics and music. At T_2 the musical composition has turned into an arranged, performed, and recorded song. While at T_1 copying the musical composition would create an identical product, at T_2 in order to get an identical product one has to duplicate the sound recording, or copy every element thereof, including the arrangement and the performers' voices. By contrast, a sound recording using the same musical composition but performed by another performer (even using the same arrangement), would yield only similar, but not the same, product. But still, at T_2 it is difficult to determine *ex ante* which of the two sound recordings is more valuable,²⁶ so despite the differences both works remain relatively close substitutes.

T₃ – From Product to Platform

At T_3 the competitive dynamics change even further. Network effects of various kinds have turned one product into a winner, to the exclusion of similar products with which it previously competed at T_2 . At T_3 the “product” is not the same product as in T_2 . It derives its value not only from the intrinsic qualities of the “work” and the additional inputs that other co-creators contributed, but also from the additional inputs added by external participants and from its ability to serve as a platform for other products and interactions to build on.

Of course, this characterization of what the product is at T_3 applies only to a minority of works, the winners. For the losers, there isn't much difference between T_2 and T_3 , except for the fact that at T_2 there was at least potential value which no longer exists in T_3 .²⁷

The difference between what the product is at T_2 and what it is at T_3 affects the relative strengths of the competitive forces to which the product will be vulnerable.

²⁶ Again, ignoring the effect of the author's or performer's existing reputation.

²⁷ This is of course oversimplification. An insignificant number of works, while not becoming superstars, may still create enough interest in smaller markets or market niches and generate, albeit on much smaller scale, the same type of network effects.

While the winner may always remain vulnerable to Schumpeterian competition, the threat of substitutive competition may come almost entirely from unauthorized identical copies; from flagrant “piracy”, but not from non-literal copying.²⁸ Returning to the previous example, even a very close cover version of a song isn’t likely to be viewed by consumers as a close substitutes and isn’t likely to undermine its market. If it does, the probable reason is creative destruction. The new version adds something new—additional value—which the first song lacked.

We can see, therefore, that across time, as the work evolves from a prototype to product and from product to platform, the relative strengths of the two competitive forces in play—substitutive competition and Schumpeterian competition—vary. Substitution effects dominate T_1 . They remain significant at T_2 but become dominated by Schumpeterian effects at T_3 . Because incentive theory of copyright seeks to maintain the incentives for creative activity more by limiting substitutive destruction but less by limiting creative destruction, it may make sense to vary the scope of copyright protection accordingly.

A few cautionary notes: New works can affect the market for existing works in more complex ways. They may reduce the demand for the existing work because they are substitutive, or they may decrease the demand for them through the process of creative destruction. But in some cases they may enhance the demand for the existing work (consider a film based on a book that leads to increased interest in the book, or a commentary that increases the demand for the original), while in other cases they may create a new market without affecting the market for the existing work (e.g., people who watch only the film but would never read the book). The fact that a new work does not affect, or even enhances, the demand for the old does not directly imply that copyright in the first work should not extend to cover the new. Arguably, extending copyright to such work increases the incentive to create the work in the first place. In fact, copyright over derivative works assumes just that. Moreover, identifying that a new work harms an old one through creative destruction instead of mere substitution also does not imply that the

²⁸ Recent empirical work on the effects of file sharing suggests that even identical copies of sound recordings are not always perfect substitutes. [expand].

creator of the new one should not have acquired a license. “Too much” Schumpeterian competition can also affect the incentives to create the work in the first place, although restricting this type of competition involves a trade-off that does not exist in the case of substitutive competition. But my purpose isn’t drawing the exact scope of copyright. My purpose is only to demonstrate that determining where the line should be drawn should consider time as a relevant factor. For this purpose I will make the simplifying assumption that copyright law seeks to encourage Schumpeterian competition, or at least is willing to tolerate it more than it tolerates merely substitutive competition.

Let us turn now to examine how these insights may bear upon several copyright rules and doctrines.

IV. THE MECHANICAL COMPULSORY LICENSE

The compulsory license regime under Section 115 of the US Copyright Act nicely reflects the differences in how others’ use of a musical work affect the work over time. It is therefore a relatively rare example of a case in which copyright law is explicitly sensitive to time within the copyright term. Section 115 allows “any person”, under certain conditions, to make and distribute sound recordings (“phonorecords”) of a nondramatic musical work when sound recordings of the musical work have been previously distributed to the public in the United States under the authority of the copyright owner, if that person notifies the copyright owner and pays a specified royalty. While the compulsory license applies to the musical composition, it does not authorize the licensee to duplicate and distribute the sound recording that contains the musical work.²⁹ In other words, the licensee must assemble “his own musicians, singers, recording engineers and equipment, etc. for the purpose of recording anew the musical work that is the subject of the compulsory license.”³⁰

Interestingly, Section 115 distinguishes between different time points that change the scope of copyright holders’ rights. The event that triggers the compulsory license is

²⁹ Other than those sound recordings made under the license. See Melville B. Nimmer & David Nimmer, Nimmer on Copyright, Release 68 §8.04[A] (2005).

³⁰ *Id.*

the distribution of sound recordings authorized by the copyright owner. That is, the compulsory license is available only when $T > T_2$ —after the completion of transition from work to product released to the market. As long as the work remains between T_1 and T_2 , the copyright owners retain full exclusive rights. It is noteworthy that under the 1909 Copyright Act, the comparable compulsory license was triggered much earlier, sometime between T_1 and T_2 , upon the making or licensing of the first recording, even if no authorized records have been distributed to the public.³¹ Congress considered the availability of the compulsory license at this period “unfair and unnecessarily burdensome to copyright owners”.³² Moreover, the compulsory license is available only if the licensee’s primary purpose is making sound recordings for distribution to the public for private use, as distinguished from sound recordings intended primarily for use by commercial users such as broadcasters, jukebox operators and background music services.³³

Historically, the compulsory license was born in 1909 out of Congress’ intention to grant to musical work copyright owners the right to control the "mechanical reproduction" of their works and overturn the Supreme Court's opinion in *White Smith v. Apollo Music*³⁴ which ruled that player piano rolls were not “copies” but rather were component parts of machines. At the same time, however, Congress sought to address the concern that one piano roll company, the Aeolian Company, would dominate the market for piano rolls.³⁵ Nevertheless, Congress’ recognition (even if implicit) of the difference in the competitive forces musical works face across time may explain the specific details of the *current* regime, which Congress redesigned in 1976. Extending a compulsory license to duplicates of the authorized sound recording would create a perfect substitute for these sound recordings and would directly harm the copyright owners because the prospect of perfect competition would restrict producers’ willingness to pay royalties higher than those set statutorily or even to pay any royalties. It therefore makes

³¹ *Id.*, at §8.04[C].

³² H.R. Rep. No. 83, 90th Cong., 1st Sess. 67 (1967).

³³ *Id.*, at §8.04[D].

³⁴ 209 U.S. 1 (1908).

³⁵ Lydia P. Loren, *Untangling the Web of Music Copyrights*, 53 CASE W. RES. L. REV. 673 (2003), at 680-81.

sense to prohibit duplicates of the sound recordings. Similarly, prior to T_2 ($T < T_2$), a sound recording made under a compulsory license also has a potential to function as a close substitute to the one authorized by the copyright owner, so similarly it makes sense not to allow a compulsory license at this stage either. However, a cover version of a song already distributed ($T > T_2$) isn't likely to be a perfect substitute to the one already on the market, especially as T approaches T_3 .

The distinction between sound recordings made primarily for distribution to the public for private use and sound recordings made primarily for commercial use also reflects similar logic. The demand for sound recordings by private users (many of which are by definition music fans) is probably less elastic than the demand by commercial users who use individual songs as relatively more fungible inputs necessary to compose attractive play lists or background music. Therefore consumers from each group would respond differently to cover versions made under a compulsory license. Commercial users might be more willing to substitute a cover version for the authentic if offered a lower price and if the cover version perfectly or closely imitates the original (perhaps jukebox operators and background music services more than broadcasters). By contrast, it is highly unlikely that music fans would be willing to make such substitution. In fact, as T approaches T_3 music fans would probably reject very close imitations, but appreciate cover versions that add their unique interpretation of the songs.

The compulsory license thus preserves full exclusive rights when the work is most vulnerable to substitution, such as at T_1 or when used primarily for commercial purposes, but allows others more easily to build upon the work after T_2 when the threat of substitutive competition wanes.

V. LICENSES, DRM, REVERSE ENGINEERING, PREEMPTION, AND COPYRIGHT MISUSE

The distinction between how time affects the types of competition affecting copyrighted works may also shed light on the controversy about copyright holders' use of licensing terms and digital rights management technologies (DRM) to supersede the

initial allocation of entitlements under the copyright act. It is unsettled yet to what extent can copyright owners prevent their customers or licensees from criticizing the work, reverse engineer it, or otherwise build upon or modify it, or prevent the copying of non-copyrightable matter. While these attempts have been sometimes attacked on grounds of preemption, antitrust or copyright misuse, with occasional sympathy of courts to such claims,³⁶ generally courts do uphold such restriction, finding such contractual terms and their technological equivalents valid, enforceable and not preempted by the federal copyright law.³⁷

Critics of such restrictive practices raise the concern that they upset the delicate balance created by the copyright act. For example, if copyright law considers fair use essential if copyright law is to serve the public interest, or allows reverse engineering under such circumstances, copyright owners should not be allowed to replace the law of the land with their own contract-made (and / or technology-backed) law.³⁸ The typical response is that contract claims are qualitatively different from copyright claims and therefore the concern simply misconceives the issue. As Judge Easterbrook explained: “[a] copyright is a right against the world. Contracts, by contrast, generally affect only their parties; strangers may do as they please, so contracts do not create ‘exclusive rights’”.³⁹ Therefore “licenses are enforceable unless their terms are objectionable on grounds applicable to contracts in general (for example, if they violate a rule of positive law, or if they are unconscionable).”⁴⁰ I do not intend to resolve this debate here. My purpose is only to illuminate how factoring in time can affect the legitimacy of such contractual/licensing/technological restriction. I suggest that earlier in time, at T_1 , there may be more legitimate reasons to uphold such restrictions which are consistent with copyright underlying policies, but which may wane further down the road.

³⁶ See,

³⁷ See e.g., *Bowers v. Baystate Techs., Inc.* 320 F.3d 1317, 1325-26 (“[P]rivate parties are free to contractually forego the limited ability to reverse engineer a software product under the exemptions of the Copyright Act” and “a state can permit parties to contract away a fair use defense or to agree not to engage in uses of copyrighted material that are permitted by the copyright law, if the contract is freely negotiated.” *id.* at 1337).

³⁸ See generally. Margaret J. Radin, *Regime Change in Intellectual Property: Superseding the Law of the State with the “Law” of the Firm*, 1 U. Ottawa L. & Tech. J. 173 (2004).

³⁹ *ProCD, Inc. v. Zeidenberg*, 86 F.3d 1447, 1454 (7th Cir.1996).

⁴⁰ *Id.* at 1449.

In describing the typical life cycle of copyrighted works, we noted how a T_1 prototype “work” transforms into a T_2 “product” and later to a T_3 “platform”. We also noted that the transition from T_1 to T_2 usually involves contribution by various co-producers, and that at T_1 the work often may be easily substituted by other works. This implies that at T_1 the author is vulnerable to opportunistic behavior of co-producers. The information the author conveyed to co-producers cannot be unlearned and the co-producers may now renege on their contractual obligations by threatening to substitute the work with a non-infringing one or simply start competing with it.

Suppose a software company hires a group of experts to run a beta version of its program, examine it, identify bugs and possible security holes and suggest improvements. The experts may find it attractive, after learning not only the benefits of the software but also its flaws, to create and market their own competing and possibly improved software. Alternatively, they may simply threaten to do so in order to re-bargain their negotiated terms. Anticipating this possibility, the software company may seek to prevent the experts from distributing the improved version of the software without its consent. Copyright law may allow the company to get an injunction against the expert if they breach and distribute copies containing its source code, but would not help if the experts write a new, non-infringing code, perhaps by reverse engineering the software. To address these concerns, the company may seek to prevent the experts from reverse-engineering its product or even from writing any competing code without its consent. The software company is also concerned that if the security holes are disclosed, writers of malicious code may take advantage of them before they are fixed, so it may require the experts not to disclose the flaws that they find. So far, such contractual restrictions do not seem to be highly objectionable, despite the fact that the contract may derogate from some of the rights to which the experts would ordinarily be entitled to under copyright law. For example, copyright law does not prevent others from writing competing but non-infringing code; with some restrictions, the law does not prohibit the experts from creating improvements or reverse engineer the software, and fair use and the Constitution certainly allow the experts to publicly criticize flaws that they discover. But unless we believe that the set of entitlements as contained in the copyright act are, and always are,

socially optimal, it is difficult to think why we wouldn't regard this set of entitlements as a default template, around which parties can contract to create the most efficient results. The distinction between the *in rem* and *in personam* effects of copyright versus contract may be useful to alleviate the concern that the contractual restrictions would stifle innovation, democratic discourse, etc. While it is true that the *in personam* characterization of contracts in this context ignores the externalities created by such contracts on third parties (i.e., the public), in the sense that the restrictions impose a negative externality on consumers who would have bought the improved product⁴¹ (or generally would have been interested in the information generated by the experts), the externality created at T_1 is relatively small. As long as only the group of experts is bound by the contract restrictions, whereas others remain free to create competing and improved program and say whatever they wish about the software, the concern that such negative externality would be serious seems remote. Therefore at T_1 such restrictions cannot be highly objectionable.

At T_3 however, a similar set of restrictions, if contained in an End User License Agreement (EULA), would justifiably look more problematic. Assuming that the software has already secured significant market share,⁴² the concern that the restrictions may hamper the process of creative destruction seems more plausible. At least, whatever legitimate concerns the restrictions are designed to address, they must be weighed against this concern. Moreover, the distinction between copyright and contract is less useful at T_3 as a much greater number of people is bound by the restrictions, and in fact breaks down completely if the restrictions are encoded into the software and circumventing them becomes illegal, as they practically bind even those who are not privy to the contract.⁴³ The minor negative externality created by the restrictions at T_1 may become a major social cost at T_3 . Even if the restrictions may serve some beneficial outcomes at T_3 ,⁴⁴ their potential so stifle creativity and discourse may justify greater suspicion.

⁴¹ See Mark A. Lemley, *Beyond preemption: The Law and Policy of Intellectual Property Licensing*, 87 Cal. L. R. 111 (1999), at 170.

⁴² Works that gain greater market share are the ones that typically attract others to build upon.

⁴³ Lemley *supra* note 41, at 148.

⁴⁴ In *ProCD* for example, Judge Easterbrook highlighted how the prohibition on the copying of facts allowed the producer of the database to price discriminate between high-value and low-value users and

Consider another example. It is commonplace in academia to distribute copies of their work-in-progress bearing a plea “Draft, please don’t cite without permission”. Ignoring for the sake of argument that legally the plea is only a request, not a binding commitment on readers,⁴⁵ I believe that even the most avid defenders of the public domain would honor such a request and would believe that others should honor it too. I also believe that even the most avid supporters of expansionist copyright would denounce a similar request if affixed to a published article or book. The difference is that the request supports creativity in the case of a work-in-progress (T_1) but hinders it after the work is published ($T \geq T_2$). Recall that review by others is an important aspect in the transformation from a work to product, but also that the value of a work often depends on what other people say about it. Therefore, at T_1 a negative review can be devastating to the work’s future success. Authors understandably seek to get comments from others (even negative ones) so that they can improve their work, but wouldn’t necessarily want the comments to become public prematurely, i.e., before they feel they have completed the work. They optimize this trade-off by disseminating their drafts widely but include a no-citation condition. If the condition cannot be honored, authors would rather disclose work-in-progress only to a small circle of peers whom they can trust. The expected result would be less pre-publishing scrutiny and possible lower quality of published works.⁴⁶ In this case the fostering creativity justifies a T_1 limitation on one of readers’ most fundamental and least controversial liberties: the right to cite and comment others’ work. By contrast, by publishing the work the author can no longer improve it. No-citation condition at this stage cannot serve the interest of improving the work, but can only

thereby to sell more copies at lower prices. Recently, in *Davidson & Assoc. v. Jung*, 422 F.3d 630 (8th Cir. 2005), the court upheld terms prohibiting reverse engineering. The prohibition there helped copyright owners in computer game software and online gaming service software to make sure that people who used pirated copies of the software would not be able to access the online gaming service. This may be a valid justification. Note however that in both cases the purportedly beneficial purpose of the restriction served only as a background reasoning, whereas the contracts were upheld simply on the basis of the distinction between contract and copyright.

⁴⁵ To overcome this difficulty, we can imagine a electronic depository of working paper such as SSRN offering authors a feature of a “clickwrap” license that allows readers to read or download the paper only if they agreed to such no-citing term.

⁴⁶ Another negative outcome is a delay in the dissemination of new ideas. Scholars benefit greatly from early exposure to cutting-edge ideas, even if they are still not fully developed or articulated.

suppress the creation of additional works. An optimal rule, therefore, could treat fair use as a default rule at T_1 , but as inalienable right from T_2 onward.

VI. THE IDEA-EXPRESSION DICHOTOMY

It is well established that copyright cannot subsist in ideas,⁴⁷ only in the specific expressions of ideas. While no one can copy the expression without permission, the ideas contained therein are free to all. Of course, not only the exact words chosen in the “expression” are protected; copyright’s scope is broader than that. As Judge Learned Hand long ago in *Nichols v. Universal Pictures Corp.*⁴⁸ copyright “cannot be limited literally to the text, else a plagiarist would escape by immaterial variations.”⁴⁹ If it were, “[t]he economic motivation of creation that underlies copyright would be almost completely vitiated if anyone could, with impunity, take an author’s work by the device of making a few changes in wording, or even by closely paraphrasing the entire work.”⁵⁰ The assumption is that copying with immaterial variations results in a work that functions as a very close substitute to the original, thus undermining the incentive to create. On the other hand, extending copyright to cover “ideas” would do disservice to the very purpose of copyright law, as it would stifle other creators’ ability to create their own work, and participate in the process of creative destruction. The idea-expression dichotomy therefore seems to reflect and serve the distinction between substitutive and Schumpeterian competition. It prevents competition that comes from perfect or near-perfect copies, but preserves the ability of built-upon works to creatively destroy existing ones.

The idea-expression dichotomy serves a beneficial purpose if its value is assessed at T_3 (or later). Realizing that at T_3 substitution effects come primarily from perfect or near-perfect copies, but less so from imperfect ones, it makes sense to protect only the specific expression of ideas but not the ideas themselves. But earlier, in T_1 or even T_2 , it is less clear that the idea-expression dichotomy serves copyright’s incentive purpose as

⁴⁷ 17 U.S.C. §102(b).

⁴⁸ *Nichols v. Universal Pictures Co.*, 45 F.2d 119 (2d Cir. 1930), *cert. denied*, 282 U.S. 962 (1931).

⁴⁹ *Id.*, at 121.

⁵⁰ ,Nimmer & Nimmer *supra* note 9, at §1.10[B][2].

good as it does at T_3 . At T_1 a work that copies another work's ideas but modifies their expression enough to fall on the safe side of the dichotomy may still function as very close substitute. At this stage both works are relatively fungible prototypes and the copy may undermine the first work's market without necessarily adding much of value.

Consider the following example. During his first year as an economics assistant professor at Berkeley George Akerlof wrote the paper "The Market for Lemons".⁵¹ By June of 1967 the paper was ready and Akerlof sent it to *The American Economic Review* for publication. Shortly he received a rejection letter in which the editor explained that the *Review* did not publish papers on subjects of such triviality.⁵² After a few other rejections on similar and other grounds, the paper was finally accepted and published by the *Quarterly Journal of Economics* in 1970.⁵³ In 2001 Akerlof shared a Nobel Prize in Economics.⁵⁴ In its decision to award the prize, the Royal Swedish Academy of Sciences explained that Akerlof's paper "is probably the single most important contribution to the literature on economics of information. This paper has all the typical features of a truly seminal piece. It introduces a simple but profound and universal idea, offers numerous interesting implications and points to broad applications."⁵⁵ This is, of course, a T_3 ex post perspective. Ex ante, at T_1 , even experts in the fields thought that the paper was trivial, just one among many other trivial papers.

While it may be true that the motivation of most academics to write and publish does not depend on copyright, many of them are driven by the prospect of winning promotion, recognition and prizes, which often depends on the whether they publish original contributions and on how widely cited those publication become. Being the first paper to come up with a new idea is important because it increases the chance that other scholars would cite this paper, and a cited paper is more authoritative than a similar but less cited one, and therefore is likely to be cited even further. So suppose that prior to its

⁵¹ George A. Akerlof, *The Market for Lemons - Quality Uncertainty and Market Mechanism*, 84 *Quarterly Journal of Economics* 488 (1970).

⁵² George Akerlof, *Writing the "The Market for 'Lemons'": A Personal and Interpretive Essay* (2003) at http://nobelprize.org/nobel_prizes/economics/articles/akerlof/.

⁵³ *Id.*

⁵⁴ Akerlof shared the Prize with economists Michael Spence and Joseph Stiglitz.

⁵⁵ Royal Swedish Academy of Sciences, *Markets with Asymmetric Information 2*, Oct. 10, 2001 available at http://nobelprize.org/nobel_prizes/economics/laureates/2001/ecoadv.pdf.

acceptance, someone else, who had read the manuscript and recognized its ingenuity, decided to write her own version of Akerlof's theory, but to express it somewhat differently. Suppose that *she* had submitted it to the *Quarterly Journal of Economics*, just before Akerlof did, so that instead of accepting his paper, the editors had decided to accept hers. If that happened Akerlof's prospects to publish his original contribution (and win a Nobel Prize) would have been frustrated, and so would the academic incentive structure.⁵⁶ Objecting to such T_1 copying of ideas therefore makes sense from an incentive perspective. Later down the road, however, protecting ideas might be more harmful for creativity than allowing their copying. Post-publication Akerlof's contribution has received the recognition it deserves (or at least had to opportunity to), and allowing others freely to build upon those ideas is desirable for all of the well known reasons.⁵⁷

Copyright law however does not make this distinction between the debilitating effect that copying the ideas would have on incentives at T_1 and the beneficial effect of copying the same ideas at T_3 . But assuming the plagiarist had submitted her paper without attribution to the manuscript she had read, she must have breached the strong norms against plagiarism in academia and consequently face the risk of both formal and informal sanction. Therefore, copyright's uniform treatment of ideas across time may not be a serious problem in academia, as the incentives at T_1 are preserved by the extra-legal norm. However, when such extra-legal norms do not exist, the uniform treatment of ideas across time may be more troublesome.

Consider the example of television formats. A writer generates a concept for a new TV series, or a "format" which may include "storylines, character descriptions, talent selection, setting, music, game rules, script treatments, production guidelines, etc. [which

⁵⁶ This is true even if in her version she addressed some of the weaknesses in the original paper. Although the result is a "better" paper, substitution effects still dominate, as most of the value lies in the original contribution of the first paper, not in the improvement.

⁵⁷ Unacknowledged copying may still cause harm at T_3 but as Landes and Posner note, at this stage the principal victims are those who credited the plagiarist and bestowed upon her benefits that she does not deserve, or people who directly competed with her for those benefits, *see* Landes & Posner *supra* note 17, at 62. The harm caused by plagiarism also depends on the genre. Readers of popular books are less interested in identifying the exact original contribution of the author than readers of professional literature, *id.*

become] a blueprint for production.”⁵⁸ She proposes the show to several interested broadcasters and enters into a contract with one broadcaster, but then another interested broadcaster who did not win the bid takes the idea and creates a similar show.⁵⁹ Copyright infringement claims generally fail as courts often find that the similarity is only in unprotected ideas or scenes a faire, but not in the expression of those ideas.⁶⁰ Moreover, on ground of preemption, courts often reject claims under state law based conversion, misappropriation and quasi-contract, leaving contract theories as the only viable cause of action against this form of plagiarism.⁶¹

In both examples, there may be marked differences in how plagiarism affects competition and the incentives to create as we move across time from T_1 to T_3 . At T_1 the value of the work lies predominantly in the ideas it conveys. Therefore, a paper that plagiarizes the ideas can function as a very close substitute. If the two compete over who would get published, the publisher has no clear reason to prefer the one over the other. Because who gets published first may determine to the benefit of whom network effect would work, plagiarism at T_1 can be devastating to the original author. The original author may find that no publisher is interested in his work anymore, or even if he does make it to T_2 that the prospect of being cited and make it to T_1 has been preempted by the plagiarist earlier publishing. The same is true for the TV format. At T_1 the broadcaster can be quite indifferent between the two similar formats. Even if both shows make it simultaneously to T_2 viewers may not necessarily have good reason to prefer the original over the copycat, and both may have equal opportunity to make it to T_3 . If making it to T_3 is the reward for investment in T_1 and T_2 the absence of tools to prevent others from preemptively making it to T_3 can adversely impact the incentive to create in the first place.

Yet once a TV format is aired, the lack of copyright protection to the format seems less problematic from an incentive perspective. Viewers would probably prefer

⁵⁸ Jay Rubin, *Television Formats: Caught in the Abyss of the Idea/Expression Dichotomy*, 16 Fordham Intell. Prop. Media & Ent. L. J. 663 (2006), n1.

⁵⁹ *Id.* at 664-65 (documenting several recent examples of this scenario).

⁶⁰ *Id.* at 670.

⁶¹ *Id.* at 668.

watching the original and wouldn't easily switch to another show which copies the format, unless the new show uses the similar ideas in some preferable way. Protecting the ideas at this stage has a weaker incentive-based justification, and may deter, or at least raise the cost of subsequent creativity. An indication that no protection of ideas is less problematic at T_3 is that a thriving international licensing market for TV formats has emerged.⁶² Although copyright law does not protect the format, the law prevents other broadcasters from broadcasting the successful show in its entirety, and may provide protection to some elements of the show, which may also be trademarked. These, as well as the short life-shelf of TV formats⁶³ and the advantage of behind-the-scenes expertise,⁶⁴ may suffice to make licensing the format more attractive than reverse engineering it for those who wish to take advantage of the success of the format without adding much of their own. This may suffice in protecting the successful show against merely substitutive competition. But at the same time, leaving the ideas free for other producers to take allows them to compete by offering something else that the first one lacked. If they are successful, it is because creative destruction worked once again.

The preceding discussion suggests that protection of ideas at T_1 may be desirable from an incentive perspective, even if undoubtedly should be rejected at T_3 . This proposition may raise several objections. The first is conceptual. After all, unlike fair use, the idea/expression dichotomy does not lend itself as easily to the same degree of flexibility. Although “[n]obody has ever been able to fix the boundary [between ideas and expressions] and nobody ever can”⁶⁵ conceptually the boundary clearly exist and if the copyright law categorically considers ideas non-copyrightable it is difficult to see how they can be, even if it is desirable to do so. But of course, the Act may be reformed, and the insights suggested here may influence courts’ approach towards state-based “law of ideas” and the question of whether such state law is preempted by the federal copyright act.

⁶² Gautam Malkani, *Television - Haven't We Seen That Programme Somewhere Before? Got Any Good Ideas? If So, Beware the Copycats, as Protection of TV Formats is Weak and You'll Need a Detailed 'Bible' to Stop the Rip-offs*, Financial Times, Sept. 21, 2004, at 8. (noting that the TV format licensing business is worth hundreds of millions of British pounds in licensing revenues).

⁶³ *Id.*

⁶⁴ *Id.*

⁶⁵ Nichols, *supra* note 48, at

Another type of objection may involve arguments about cost of error. Assuming that the idea-expression dichotomy serves well the purpose of copyright at T_3 but less so at T_1 the question is whether we can tailor a rule that will allow appropriate protection at T_1 and deny it at T_3 or whether trying to craft such a rule would inevitably lead to T_3 copyright owners successfully fending off legitimate competitors by disguising themselves as T_1 victims, and T_1 copiers successfully passing themselves off as innocent T_3 borrowers of ideas. Another contingency that should be considered is that allowing T_1 claims for idea protection would increase the number of nuisance plaintiffs harassing successful copyright owners claiming that their ideas were stolen. Even now many big content creators refuse to accept non-solicited ideas for fear of litigation,⁶⁶ and it can be expected that this inefficiency (after all, there must be some good ideas out there) would grow if suing becomes easier.⁶⁷ If we cannot avoid or minimize these costs, then perhaps we are better off with the current rule, which assures that ideas remain in the public domain, even if occasionally the result is some disincentives at T_1 . We saw that at least in academia the extra-legal norms against plagiarism address the problem,⁶⁸ so it may be useful to know the extent of the problem in other areas such as TV formats before upsetting the idea-expression dichotomy can be considered desirable.⁶⁹ In fact, as Christopher Fay, the managing director of the German-based Format Recognition and Protection Association noted: “TV lives from borrowing from what has gone before ... The worst thing would be for a judge to make the wrong decision, such as granting a monopoly on chat shows.”⁷⁰

⁶⁶ See e.g., *Preston v. Century Fox Canada* ____.

⁶⁷ This concern may be addressed at the remedy stage. What often motivates such lawsuits is the ability to get an injunction against the production or distribution of the work and the use of the threat of injunction to hold up the producer and get a settlement worse much more than the ex ante worth of the appropriated ideas (or expression). The Supreme Court’s recent decision in *eBay v. MercExchange* ___, which departs from the previous automatic issuance of injunctions in patent (and copyright) cases may help preventing this type of opportunistic litigation.

⁶⁸ Arguably, such norms seem to be a better solution than copyright within academia. The vision of academics suing each other for copyright infringement is quite disturbing.

⁶⁹

⁷⁰ Malkani, Gautam *supra* note 62.

VII. FAIR USE

[to be completed]

VIII. CONCLUSION

[to be completed]