Volume Title: Innovation Policy and the Economy, Volume 9

Volume Author/Editor: Josh Lerner and Scott Stern, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-40071-9

Volume URL: http://www.nber.org/books/lern08-1

Conference Date: April 15, 2008

Publication Date: February 2009

Chapter Title: Copyright as Innovation Policy: Google Book Search from a Law and Economics Perspective

Chapter Author: Douglas Lichtman

Chapter URL: http://www.nber.org/chapters/c8184

Chapter pages in book: (55 - 77)

3

Copyright as Innovation Policy: Google Book Search from a Law and Economics Perspective

Douglas Lichtman, *University of California, Los Angeles*

Executive Summary

The copyright system has long been understood to play a critical role when it comes to the development and distribution of creative work. Copyright serves a second fundamental purpose, however: it encourages the development and distribution of related technologies such as hardware that might be used to duplicate creative work and software that can manipulate it. When it comes to issues of online infringement, then, copyright policy serves two goals, not one: protect the incentives copyright has long served to provide authors and at the same time facilitate the continued emergence of innovative Internet services and equipment. In this chapter, I use the Google Book Search litigation as a lens through which to study copyright law's efforts to serve these two sometimes-competing masters. The Google case is an ideal lens for this purpose because both the technology implications and the authorship implications are apparent. With respect to the technology, Google tells us that the only way for it to build its Book Search engine is to have copyright law excuse the infringement that today by design is part of the project. With respect to authorship, copyright owners are resisting that result for fear that the infringement here could significantly erode both author control and author profitability over the long run. I myself am optimistic that copyright law can and will balance these valid concerns. The chapter explains how, discussing not only the formal legal rules but also the economic intuitions behind them.

On the surface, the copyright system is a set of legal rules designed to encourage the development and distribution of creative works such as books, movies, and motion pictures. Dig a little deeper, however, and it becomes immediately clear that copyright law has substantial implications for another type of innovation, namely, innovation with respect to related technologies such as hardware that can be used to distribute copyrighted work and software that can be used to manipulate it. As a result, when courts and lawmakers decide the extent to which, for example, Internet service providers will be liable for their role in facilitating

978-0-226-40071-6/2009/2009-0003$10.00

infringement online, copyright law is actually altering the incentives faced by two important groups: the authors whose work might be protected by these sorts of interventions and who have long been understood to fall within copyright's purview, and the technologists who will respond to copyright liability by altering their patterns of research, deployment, and investment.

The central challenge facing the modern copyright system is the challenge of how best to account simultaneously for these two groups. That job is not easy. Within the copyright community, there are many who are unwilling to concede the importance of building into the law safeguards designed to minimize the uncertainty and financial burdens that copyright might otherwise impose on technologists. Outside the copyright community, meanwhile, there is a striking tendency to overstate those costs while almost completely ignoring the important role copyright has long played in terms of encouraging authorship.

Into that environment comes now litigation over the Google Book Search project. Those who view copyright as an inefficient constraint on technological advancement hold out this litigation as a poster child. The Google project, after all, promises this amazing resource through which all of us would be able search the world's books in much the same way that Google today allows us to search the Web. But if copyright law obligates Google to ask permission before including copyrighted work in the database and as a result Google ends up forced to share proceeds with or make other concessions to authors, Google will suffer increased costs and might have less of an incentive to pursue this or similar projects. On the other side, the traditional copyright community rightly wonders what makes that argument special. Motion picture studios would prefer a world in which they could use preexisting books and plays without paying a license fee or in other ways negotiating permission from the relevant copyright holders, and yet no one seems to think that copyright ought to step aside in that situation and give the new technology a free pass at the expense of the old.

The legal lens through which this fight will be judged is the doctrine of fair use. I will detail that doctrine more fully later in the chapter, but at a high level the fair use doctrine is a flexible policy tool that allows courts to waive off copyright infringement in instances in which the costs of protection seem to outweigh the benefits. One way to think of fair use is to recognize that there are a large number of rights and revenues that could plausibly be assigned to authors; so if the end goal of the copyright system is only to move a certain amount of value anyway, copyright can and should choose the subset of those rights and

revenues that will transfer whatever value is necessary from an incentive perspective but do so at the lowest external cost in terms of avoidable, undesirable third-party implications.

My purpose in this chapter, then, is to use Google Book Search specifically and the fair use doctrine more generally to show how copyright law can, should, and in fact does efficiently work to achieve the dual goals of supporting authorship on the one hand and supporting technological innovation on the other. I do not myself believe that all copyright law strikes this balance correctly. The legal rules that govern third-party liability, for example, seem in some instances to give technologists far more freedom than would be efficient[1] and in other instances to punish too sharply what might be inevitable, good-faith mistakes.[2] The fair use doctrine, however, works. This is not to say that every application of the doctrine achieves the optimal, efficient result. Indeed, it is impossible to prove that any specific answer is optimal because that analysis turns on unknowable questions about exactly how much innovation and authorship would occur under alternative legal regimes and how valuable specific innovation and authorship contributions are in the first place. But fair use opens the door for the law to do its core job: recognizing all the relevant trade-offs and variables and then armchairing a plausibly efficient outcome despite the obvious information constraints.

I divide the chapter into four sections. In Section I, I briefly introduce the facts of the case. In Section II, I summarize the core legal arguments and policy intuitions. In Section III, I delve into the details, discussing in some detail the fair use doctrine and the many considerations that inform it. In Section IV, I offer a brief conclusion.

I. The Facts

Google is in the process of creating an online search engine that allows users to search the full text of published books. To use the search engine, users enter a search term or phrase, and Google's computers then look for books that might use the term or phrase and hence might be of interest. The books about which there is controversy are those that Google obtains from various libraries. The libraries allow Google to borrow books from their collections, to scan those books into electronic form, and ultimately to include the resulting electronic information in whatever databases Google builds in order to run its search service. The libraries do not hold copyright in the books, and thus the libraries themselves have no power (from a copyright perspective) to authorize Google's use.

Google scans the books it borrows in their entirety, and it stores all that information in a way that allows Google to respond to any search query that might be submitted in the future. Thus, presumably, Google saves all or most of the text of every book in some sort of database. Users of Google Book Search, however, do not see the full text of a book unless the relevant copyright holder has given permission. Instead, Google returns what it describes as "snippets," which seem to be excerpts that run only a few sentences long and contain the desired search terms. These excerpts in theory show enough information that users can evaluate whether a given book is indeed of interest. Google has proprietary software that is designed to ensure that users cannot see too many excerpts from the same book, for example, through repeated searching.

Google has publicly committed to leave certain books out of its database, including thesauruses and anthologies of short poems. It unilaterally decides which books to leave out, but the idea is to exclude books for which most of their value comes from having the ability to access a small, relevant excerpt at the right time. Google has not published a list of the books excluded, nor has it made public the details of how it selects these titles. It also allows copyright holders to "opt out" of the Google Book Search program. Specifically, a copyright holder can notify Google that it would prefer to have a specific work removed from the database. Google presumably complies with these requests.

There are a number of services that compete with the Google service. Amazon, for example, has implemented and announced a variety of search-inside-the-book programs, including a voluntary program through which copyright holders can allow would-be customers to "look inside" a book prior to buying it and an announced program that would (among other things) allow users to electronically search participating books after they have purchased the relevant book in paper form. The book publisher HarperCollins is also experimenting with electronic delivery. Even Google itself has launched a competing service—one that waits for permission from copyright holders, but upon receiving permission reports back larger excerpts. Many other services and products similarly either are available today or are in various stages of negotiation and development.

II. The Case

Litigation is already under way over the Google Book Search project (Authors Guild 2005). The result of the case will ultimately turn on the court's interpretation of section 107 of the Copyright Act (15 U.S.C.

§107 2000). Section 107 empowers a court to excuse, on public policy grounds, acts that would otherwise be deemed to impermissibly infringe a copyright holder's exclusive rights.

Courts are required to consider four specific statutory factors when evaluating a fair use claim; however, courts are empowered to go beyond those factors and engage in a broader public policy analysis as appropriate.[3] In the end, the idea is for courts to excuse infringement in instances in which a "rigid application of the copyright statute … would stifle the very creativity which that law is designed to foster" (*Stewart*, 236; quoting *Iowa State University Research Foundation, Inc. v. American Broadcasting Companies*, 621 F.2d 57, 60 [2d Cir. 1980]). The fair use doctrine is thus enormously flexible, and by necessity it vests considerable discretion in each court.

A common misconception is that the fair use doctrine excuses any infringing use that is socially valuable. That is a clear mistake. The litigation involving Michigan Document Services provides a helpful example (*Princeton University Press v. Michigan Document Services, Inc.*, 99 F.3d 1381 [6th Cir. 1996] [en banc], *cert. denied*, 520 U.S. 1156 [1997]). The infringing products in that dispute were packets of photocopied materials. The packets were made up of excerpts from articles and books, those excerpts having been chosen by university professors for use in their specific university classes. The accused infringer was the copy center that duplicated the excerpts and ultimately sold those packets to students.

Clearly, the infringing products were socially attractive. They were products that facilitated classroom teaching, and they were produced at the direction of university faculty. Yet, the copy center that produced the packets was found guilty of copyright infringement and specifically had its fair use defense rejected (*Michigan Document Services*, 1385–90, discussing fair use).

Why was the copy center denied the protection of the fair use doctrine? Because fair use is not an inquiry into whether the accused use is valuable. Instead, it is an inquiry into whether the owner of the infringed copyright should have influence over when and how the accused use takes place. To deny fair use in the Michigan Document Services dispute, then, was not to in any way speak ill of the infringing products at issue. Photocopied university materials are tremendously worthwhile products, and no one disputes that fact. To deny fair use was instead to decide that these beneficial but infringing products ought to fall under copyright holders' sphere of influence, with the relevant copyright holders having the right to influence who produces

the packets, under what terms, and how much everyone profits from that interaction.[4]

Two intuitive considerations guided the court in Michigan Document Services and indeed more generally seem to helpfully frame fair use analysis. The first of these intuitive considerations is the degree to which a finding of fair use would undermine the incentives copyright law endeavors to create. Copyright law in general recognizes rights in authors in order to motivate them to create, disseminate, and in other ways develop their work.[5] Fair use is unattractive to the extent that it interferes with that goal. Put differently, the issue here is whether repeated findings of fair use in a category such as the one at issue over the long run would reduce author motivation to do things such as create their work, share their work publicly, and search for new related projects (*Sony*, 448). If so, on this ground fair use is unattractive since it undermines the very incentives copyright law endeavors to create.

In *Michigan Document Services*, this first consideration clearly cut against fair use. Works that would be included in university course packets would often also be works whose primary audience would be the university audience. If the authors of these works could not profit from their use in class, it was not clear from where profit would otherwise come. The prospect of fair use, then, in this case came with an obvious impact on the core incentives copyright law attempts to create.

The second intuitive consideration relevant to fair use analysis is the degree to which uses such as the one at issue would continue even without the protection of fair use. In the Michigan Document Services example, there were two plausible concerns along these lines: transaction costs could make it too expensive to acquire copyright permission even if authors were otherwise willing to license this sort of use;[6] and high licensing costs could end up leading to an inefficient underuse of copyrighted work for course materials. Neither concern ended up resonating with the court. Transaction costs seemed likely to be adequately addressed by licensing intermediaries such as the Copyright Clearance Center;[7] these entities reduce costs by offering licensees one-stop shopping for a large number of titles and offering licensors a convenient way to approach and collect from a large number of would-be licensees. Licensing rates, meanwhile, seemed adequately constrained by market competition. After all, no specific author has much market power vis-à-vis academic users because a faculty member can always assign different readings if the originally chosen work is available only at an unreasonable price or subject to unreasonable terms.[8]

Return now to Google Book Search. To the extent that Google invokes fair use to defend the entire Google Book Search program, that defense seems to fail. With respect to the first intuitive consideration, a finding that Google Book Search is fair use would clearly hurt authors. For instance, Google's scanning and storage activities expose authors to an increased risk that their works will leak out in pirated form, and Google's project more generally undermines an author's incentives to implement and profit from comparable or competing offerings.

Moreover, with respect to the second intuitive consideration, a finding of fair use is not critical in terms of facilitating the creation of the Google search engine because a great deal of the project could be accomplished through negotiated, consensual transactions. Publishers, for example, could act as helpful intermediaries, negotiating terms with Google on behalf of all the authors still under contract with the publisher. And even individual authors could opt in to the program, for instance, if Google were to create a Web site where interested authors could agree to participate or even could upload electronic copies of their work. Similarly, there is no reason to believe that licensing rates would be inefficiently high. Google can build a tremendously useful resource even if at the start it has only 30% of the world's books. That is important because it means that no single author has significant market power vis-à-vis Google; an author who demands a disproportionate share of the project's profit or undue involvement in the project's design can simply be left out of the database until that author makes a more reasonable offer.

Were Google to concede infringement for many of the works at issue but invoke fair use only to more narrowly excuse its use of books in instances in which the costs of identifying the relevant copyright holder are prohibitive, however, Google's claim would be strong. It is enormously difficult to acquire permission with respect to books that are significantly old or books for which the current ownership of rights is hopelessly unclear. As applied to that class of work, Google might be right that the only way to use those books is to invoke fair use.[9] Google could also fairly point out that the harm to that subclass of authors is small because authors who are so difficult to identify are likely also not authors who are actively profiting from or marketing their work. The main weakness with this argument is that Google in practice makes no effort to distinguish these "orphan" works from the many works for which permission would be practical. A court might require Google to undertake reasonable efforts along these lines as a condition of any fair use finding.

In summary, then, Google's fair use claim fails in its current form because its legal argument and its actual practices both sweep too broadly. Google adopts a legal position that maximally facilitates the creation of its new technology service, but that position disregards two important considerations: the law's equally legitimate interest in rewarding and protecting authors and the reality that a careful enforcement of copyright law would not much undermine Google's ability to develop the Google Book Search project.

III. The Details

Fair use is an affirmative defense to a charge of copyright infringement. Its purpose is to permit "courts to avoid rigid application of the copyright statute when … it would stifle the very creativity which that law is designed to foster" (*Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 577 [1994]; citations omitted). Fair use began as a flexible, judge-made doctrine. When federal copyright law was revised in 1976, however, fair use was codified in the statute at section 107. That codification was explicitly intended to restate the then-existing law and not to expand or contract fair use in any way.[10] Thus, even today, fair use retains the flexibility and comprehensiveness of an equitable doctrine.

The statutory provision that codifies fair use begins with a list of examples, stating specifically that "reproduction … for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research" is excused (sec. 107). The provision then goes on to identify four factors that must be considered when evaluating a claim of fair use. Those factors are as follows:

1. the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;

2. the nature of the copyrighted work;

3. the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and

4. the effect of the use upon the potential market for or value of the copyrighted work.

These factors are not exhaustive. Thus courts can and do consider other factors when conducting a fair use inquiry, emphasizing facts that might not fit within the normal rubric but still seem important to understand the dispute at hand.[11] Moreover, when considering the four explicit factors, courts do not merely count them up. Instead, courts combine

these factors with other relevant information and conduct an appropriately flexible, case-specific policy analysis.[12] "The ultimate test of fair use … is whether the copyright law's goal of promoting the Progress of Science and useful Arts would be better served by allowing the use than by preventing it" (*Castle Rock*, 141; citations and quotations omitted).

Courts typically organize their fair use analysis by first considering each of the statutory factors and then, as needed, turning to other considerations. I adopt that same framework here and discuss each of the four statutory factors, apply them to the facts at hand, and then consider issues that do not fit well under those four headings.

A. The Purpose and Character of the Use

The first fair use factor is the purpose and character of the use. One issue typically raised with respect to this factor is whether the use is commercial. The intuition is that a profit-generating user can, and thus should, absorb the costs of complying with copyright law and compensating the original author.[13]

There was a time when this consideration was significantly influential. In *Sony v. Universal City Studios*, for instance, the Supreme Court stated that "every commercial use of copyrighted material is presumptively an unfair exploitation of the monopoly privilege that belongs to the owner of the copyright" (451). More recently, however, the court has backed away from this strong stance, holding instead that "the commercial or nonprofit educational nature of a work is not conclusive" and is only one factor "to be weighed along with others in fair use decisions" (*Campbell*, 585).

The reason for this hesitation is simple: many commercial uses are at the same time strong candidates for fair use. Newspapers and television stations, for instance, are clearly for-profit entities engaged in for-profit uses. Yet, to the extent that they commit copyright infringement, they typically do so in support of the news reporting and commentary functions that are explicitly endorsed in section 107.[14] The fact that an entity has a profit motive, then, turns out not to be particularly helpful in terms of distinguishing attractive from unattractive fair use cases. At best, the commercial nature of a use serves as a weak signal that the infringer has resources that could be used to reward or empower the original copyright holder and that a requirement to do so would not substantially reduce the availability of the work in question.

A second and more important issue considered as part of the first factor is the question of whether the accused use is "transformative"

in nature. A use is transformative if it is substantially different from the original work in terms of its purpose, meaning, or effect (see *Perfect 10, Inc. v. Amazon.com, Inc.*, 487 F.3d 701, 720 [9th Cir. 2007]; *Campbell*, 579). According to *Campbell*, a transformative work does not merely supersede the original work. It is instead a work that has new features or brings new value.

Whether a work is transformative is important for two reasons. First, all else held equal, a transformative work is less likely to hurt the original author. If an infringing work has the same purpose, meaning, or effect as the original work, the infringing work likely will displace sales of the original. If the infringing work is sharply different along these dimensions, by contrast, sales could remain intact.[15]

The second reason why it is important to consider whether a work is transformative is that a transformative work brings something valuable to society. The work is not merely redundant to that which society already had. It is new and has new meaning. The fact that a work is transformative, then, makes a finding of fair use marginally more attractive. Put differently, there is little reason to trump a copyright holder's exclusive rights if the only payoff is that society would get another work that is largely indistinguishable from the original one. By contrast, if society is at least getting something sufficiently new, there might be a case for a fair use finding because getting something new is itself an attractive outcome.[16]

When all of this is applied to the Google Book Search project, the commercial nature of the use is straightforward: Google clearly is a for-profit entity engaged in a profit-motivated use designed to promote the company's long-run financial interest. Indeed, if Google were spending this much money and not anticipating an ultimate return on the investment, its management team would likely be violating its fiduciary duty to the company's stockholders. The fact that Google is not at the moment explicitly cashing in on the infringing product is of little importance. Clearly, over the long run, Google will monetize its new search engine, perhaps by introducing advertisements, by demanding a royalty on downstream book sales, or by using this new search capability to further distinguish the Google family of products from rival products offered by firms such as Microsoft and Yahoo.[17]

With respect to the transformative nature of the work, however, Google has a strong case that Google Book Search is transformative. The overall purpose of its infringement is to create a new and useful tool for locating information. I do not think that tells us much about whether a finding of fair use hurts author incentives. But it does tell us that there

is at least something to be gained by a finding of fair use. If protected by fair use, Google would put into the world a product that is both socially valuable and meaningfully distinct from the works that are being infringed. In my view, that suffices to establish that the use is transformative.[18]

B. The Nature of the Copyrighted Work

The second explicit fair use factor is the nature of the copyrighted work in question. Under this factor, courts consider the creativity of the original work. If the original work falls into a highly creative category, such as fictional novels, fair use is deemed less compelling. If the original work falls on the less creative side of the spectrum, such as a biography, fair use is deemed more appropriate. The explanation is that "some works are closer to the core of intended copyright protection than others" (*Campbell*, 586). Put differently, on this view, copyright law is primarily concerned with the protection of creative, expressive work, and as a result fair use is less objectionable when it reduces the protection given to works that are not significantly creative or expressive.

A second consideration sometimes included in a discussion of this fair use factor is the question of whether the original work is sufficiently available to the public. A work that is out of print, for example, on this argument might be more vulnerable to a fair use defense (S. Rep. no. 94-473, at 64 [1965]). The intuition here is twofold: first, fair use might be the only way to facilitate use of an otherwise unavailable work; second, a finding of fair use might not much undermine author incentives in a situation in which the author himself has already stopped promoting or otherwise offering his work to potential licensees.

Application of this unavailability concern is complicated, however, and courts have varied in their approach. In *Basic Books v. Kinko's Graphics* (758 F. Supp. 1522, 1533 [S.D.N.Y. 1991]), the court noted that it might be more important to deny fair use as applied to out-of-print works because the royalties at issue in the litigation "may be the only income" the relevant authors will earn. In *Princeton University Press v. Michigan Document Services* and separately in *American Geophysical Union v. Texaco*, two courts recognized that by denying fair use, copyright law can support the development of intermediaries such as the Copyright Clearance Center that facilitate licensing and in that way make more work accessible.[19] The influential Nimmer treatise, meanwhile, makes a related point: an out-of-print work will come back into print whenever demand is high enough and costs are low enough, but

those conditions might "never arise if competitors may freely copy the out-of-print work" (2008, sec. 13.05[A][2][a]).

Applying all this to Google Book Search produces a mixed result. Some of the infringement that takes place as part of the project would likely be favored under the second fair use factor, either because the books being infringed are more informational than creative or because the books are out of print and/or otherwise inaccessible for licensing. However, to the extent that Google scans books that are largely creative or to the extent that it scans books that are in fact available for consensual licensing, the second fair use factor would likely favor the copyright holders.

Interestingly, Google does not separate books along these dimensions when it engages in its infringing activities. It could. Google's partner libraries surely sort their collections in ways that distinguish novels from biographies. And it would be easy for Google to check, prior to scanning, whether a given book is in print or is readily available for licensing through its author, publisher, or a licensing intermediary. This failure on Google's part might be deemed to forfeit its otherwise legitimate claim to a partial victory under factor 2.

C. The Amount and Substantiality of the Portion Used

The third explicit fair use factor is the amount and substantiality of the portion used. As a general rule, the more the infringer takes, the more this factor weighs against a finding of fair use. The intuition is the obvious one: the extent of the copying is a good proxy for the harm imposed on the copyright holder. If an infringer takes only a tiny segment of a copyrighted work, the odds are low that the taking will much undermine the author's ability to exploit his own full contribution. If the infringer takes the bulk of the work, the opposite logic applies. In this sense, this third factor in some ways echoes the considerations raised under the first factor's test for transformative use and the fourth factor's test for the economic significance of the copying.

There are exceptions to the general rule stated above. For instance, copying a small amount from the original work might still be problematic under this factor if what was taken turns out to be "essentially the heart" of the work. In a famous case along these lines (*Harper & Row*), a magazine purloined a tiny portion of an unpublished manuscript, but still the third fair use factor was deemed to favor the copyright holder because the copied words represented the excerpts that would-be readers were likely most interested in seeing.[20] Conversely,

copying the entire work might not weigh against fair use in a case in which the only way to accomplish the infringing use is to copy at that scale. In *Sega v. Accolade, Inc.* (977 F.2d 1510 [9th Cir. 1992]), for instance, the infringer copied the entirety of a software program in order to study how certain aspects worked. The court put "very little weight" on the amount of copying, however, both because the complete copy was not actually used after the learning was complete and because there was no reasonable alternative means by which to dissect the program anyway (1526–27).

The Google Book Search project obviously involves the scanning of entire books, and thus to some degree the third factor will weigh against a finding of fair use. This is appropriate because it is the existence of these full copies that leads to one of the harms that most concern copyright holders: full copies might accidentally leak out. That distinguishes the aforementioned cases in which copying of the full work was excused. In those cases, full copies were made, but there was never much risk that those full copies would fall into the hands of unrelated parties. Here, the risk is significantly more pronounced.

In the opposite direction, however, note that while a workable search engine could be built through a process that used less than the full text of the relevant books, the charm of the Google project is that its search engine can search any word or phrase in the book. That is what makes Google's search index better than conventional alternatives. There are many indexes that sort books on the basis of keywords or other organization themes that are chosen ahead of time by the organizing party. Google's index is unique in that it allows the user to dynamically define the keywords that will then be used to retroactively sort the books. That feature could not be achieved without Google having access to the full text of the works.

The third fair use factor is similarly complicated as it applies to the snippets that Google offers to its users. Snippets in this context are in a very real sense the heart of each work. They are chosen on the basis of the user's own search terms, and they are designed to show the user the exact part of the book that he is most interested in seeing. Thus there is a very real analogy to be drawn to the *Harper & Row* case discussed above. In both this case and that one, the size of the infringement is not a good proxy for its economic or artistic significance; the takings in both situations are small but tremendously well targeted.

Putting all of that together, I doubt that the third factor should or will much move a court's analysis one way or the other. As I suggest above, the third factor is largely redundant to the analysis conducted under the

first and fourth statutory factors. I suspect that the third factor will therefore not be paid much attention. In this case, the other two are much more helpful in terms of sharpening the core public policy issues at stake.

D. The Effect on the Plaintiff's Potential Market

The fourth explicit factor listed in section 107 is the effect on the potential market for, or value of, the copyrighted work. This is relevant because a use that interferes with the value of the original work likely undermines the incentives that copyright law is designed to create in the first place. That is, the original idea behind copyright law was to encourage authors to create, disseminate, and in other ways promote their work by promising authors certain exclusive rights. The more a fair use finding would reduce the value of those exclusive rights, the more disruptive that fair use is to the copyright system and hence the less attractive the fair use defense.

When evaluating the fourth factor, courts consider "not only the extent of market harm caused by the particular actions of the alleged infringer, but also whether unrestricted and widespread conduct of the sort engaged in by the defendant would result in a substantially adverse impact on the potential market for the original" (*Campbell*, 590; citations and quotations omitted). That is, the fourth factor does not merely look to see whether this infringer, through its actions alone, would substantially impose author harm. The factor more broadly considers whether actions in this category, if repeated by a large number of unrelated infringers, would cause substantial author harm.

That harm, meanwhile, includes harm to "potential" markets. Thus the fourth factor is implicated not merely when the infringing use might reduce sales of the original work in its current form, but more generally when the infringing use might interfere with future exploitation of the work in other forms.[21] According to *Campbell*, relevant markets under the fourth fair use factor include markets that the author has not yet entered. One influential line of cases holds that any market can count as long as it is a "traditional, reasonable, or likely to be developed" market (*American Geophysical Union*, 930).[22]

Courts and scholars sometimes worry that this fourth factor is circular (*American Geophysical Union*, 930, n. 17). After all, if fair use is denied in a given case, then the infringer in that case would himself likely pay the author some sum in exchange for the right to continue the infringement. Can that potential payment really count under factor 4, the

result being that in almost every case factor 4 would, at least to a small degree, weigh against a finding of fair use?[23]

The answer is that factor 4 actually should in every dispute weigh at least slightly against a finding of fair use. This is not to say that fair use should be denied in every case. Instead, my point is that, in almost every case, fair use does reduce author incentives. Other considerations might then swamp that concern; but factor 4 is designed to highlight the degree to which a finding of fair use would hurt authors, and framed that way, there is no reason to exclude from the calculus the losses associated with the very use being litigated.[24]

On the facts of Google Book Search, the fourth factor weighs strongly against a finding of fair use because there are at least four types of cognizable harm. First and most obviously, Google imposes a substantial harm on authors when it scans, transmits, and stores complete electronic copies of previously nonelectronic books. The harm here comes in the form of a security risk. Google's electronic copies could leak out not only during the initial scanning process but also later in time, when the electronic copies are stored indefinitely in Internet-accessible databases. Google surely has security precautions in place to prevent the electronic versions from leaking out. However, there is no reason to believe that its security precautions are appropriate.[25] Put differently, copyright holders are harmed here because electronic duplication introduces new and substantial risks, and yet Google's project allows copyright holders no say over how those risks should be managed or what should happen in the event the risks mature into a substantial security breach.

Google would likely respond by suggesting that courts can evaluate Google's security precautions and make any fair use finding contingent on a showing of adequate security. That is in part an attractive middle ground. Analysis of factor 4, however, cautions against that approach. After all, the question here is whether "unrestricted and widespread conduct of the sort engaged in by the defendant" would impose substantial author harm. In my view, if authors are told that anyone can scan, transmit, and store full copies of their books for use in index-like products and that the only protection is after-the-fact judicial evaluation of the relevant infringer's security precautions, I suspect that authors will rightly expect that their work will leak out. Courts are just too slow and too far removed from technical details to meaningfully regulate security issues of the sort implicated here.

Second, for at least some of the works being copied, Google's act of providing snippet access will directly undermine the market for the original works. A technical dictionary, a thesaurus, an anthology of short

poems, and a book of famous quotations are each valuable in large part because users are at any given time interested in only a specific short excerpt. If Google provides those very excerpts via its online search engine, the value of these books will be sharply reduced.

As I mentioned earlier, Google itself has acknowledged this and made a public commitment not to provide even snippet access to these sorts of works. As with the security issue, however, that solution is unsatisfying both because Google's judgment might inefficiently favor its own interests and because again the proper analysis here is to consider not merely whether authors would be harmed if forced to trust Google on this matter but more generally whether authors would be harmed if snippet access of this sort were to become a widespread practice, run by possibly trustworthy firms such as Google but also by a wide range of actors with varying degrees of honorable motivation.

Third, Google's project directly undermines author opportunities to pursue projects that are similar to and/or partially competitive with Google Book Search. For instance, both Amazon and the publisher HarperCollins have announced their own services that would include electronic book access and/or book search capabilities. If Google is allowed to compete with those services under the protection of fair use, authors will have a harder time earning profits from and otherwise being successful with these other programs.

Fourth and finally, there is the purely circular harm: if Google's fair use defense is rejected, it will surely take steps to include authors in the design of the book search project and also to include authors in some of the financial gains the service makes possible. As I note above, this circular harm is a controversial consideration, but in my view the circular harm is rightly included in the factor 4 calculus. Again, the question under factor 4 is the degree to which a finding of fair use would limit author control and author profit, thereby undermining author incentives. Google's refusal to include authors in the decision-making process and its decision to deny authors any share of Google's revenues is therefore plainly relevant. If Google Book Search is even half as successful and socially important as its proponents predict, the royalties at issue in this case alone could significantly increase author incentives to write, disseminate, and otherwise invest in their work.

E. Additional Considerations

The four statutory factors play a central role in almost any fair use analysis. However, fair use also welcomes consideration of other relevant

public policy issues. Here, then, I briefly consider two issues that the parties might raise along these lines.

1. Google Book Search Benefits Authors

The popular commentary on Google Book Search emphasizes the fact that Google's search engine will likely increase demand for books. That argument resonates. By making it easier for people to identify books that might be of interest, a comprehensive search engine should in the aggregate increase book demand. This should be especially true for books that serve a niche market. Those books are hard to find in conventional ways because they are not sufficiently known or advertised, but Google's content-based search engine should compensate for those limitations, increasing the likelihood that interested readers will find these niche offerings.

That said, the fact that the Google project might in one way benefit copyright holders does not significantly change the overall fair use analysis. After all, this fact tells us only that authors are better off in a world in which Google's project is fair use as compared to a world in which no one builds book search engines at all. That, however, is not the relevant comparison. Instead, the fourth factor of the fair use inquiry asks about the degree to which authors are worse off in a world in which fair use takes away their ability to license the use or pursue it themselves. Clearly, authors would be better off if they could negotiate their own deal with Google or pursue their own versions of the search technology rather than merely receive whatever benefit the project happens to offer them by default.

None of this should be surprising. All sorts of infringing work benefits authors, and yet authors nevertheless routinely keep their right to say no. Movies that are based on books, for example, typically increase demand for the underlying books. Still, there is no question that the people who produce those movies must seek permission from, and negotiate financial details with, the relevant copyright holders. The reasons are the very ones I have considered at length here: author incentives are at stake in the question of whether or not a movie should fall under the copyright holder's sphere of influence; and because movies really do create value that can be shared by both the movie producer and the relevant book author, it seems likely that movies will still be made even if fair use is denied.

2. Google Allows Copyright Holders to Opt Out

The popular commentary also has been taken with the argument that Google's use should be deemed fair because it allows copyright holders

to opt out of the program. Specifically, a relevant copyright holder can notify Google that he does not want a particular book included in the database, and Google has promised to respect that request.

This opt-out offer certainly makes the Google project more attractive than it would otherwise be, but again my suspicion is that this feature will not and should not significantly influence the overall analysis. The reason is the fundamental insight that fair use considers "not only the extent of market harm caused by the particular actions of the alleged infringer, but also whether unrestricted and widespread conduct of the sort engaged in by the defendant would result in a substantially adverse impact on the potential market for the original" (*Campbell*, 590). An ability to opt out works well in a world in which Google is the only infringer. In that case, at low cost authors could find out about the Google project and communicate their desire to be left out if need be. This would be efficient, in fact, because the costs to authors in finding Google would likely be much smaller than the costs Google would incur were it required to find each individual copyright holder.

When the analysis shifts to focus on the possibility of countless Google-like opt-out programs, however, the conclusions reverse. In a world with a large and ever-changing list of opt-out projects, authors would be forced to invest substantial sums finding each project and notifying each about their desire to participate. The problem would be even worse if some of those opt-out programs were designed strategically to make things difficult on authors, for instance, imposing high standards of proof before acknowledging that an opt-out really came from the correct copyright holder. (Infringers have an incentive to do just that because in an opt-out system, infringers benefit if authors find it too expensive to actually engage in the mechanism of opting out.)

Overall, then, the problem with an opt-out program is that it does not scale. This is one reason why copyright more generally is defined as a permission-based, opt-in system. The opt-in approach gives copyright holders meaningful control over potential infringements. The opt-out approach, by contrast, is an expensive proposition that would substantially erode the value of copyright rights.[26]

IV. Conclusion

The fight over Google Book Search is important on its own terms. However, as I emphasized at the start, my interest here is broader.

The Google fight provides a specific example of what I expect to be the central challenge facing copyright law over many years to come: the difficulty of balancing copyright's role in encouraging authors with its possibly unintentional but also unavoidable role in influencing the development of related technologies.

My analysis of the law as it applies to Google's project suggests that copyright law can meet this challenge. New technologies do not require a complete free pass on all copyright liability in order to thrive. Just the same, authors can be efficiently rewarded even without absolute control over the use of their work. The key to getting the analysis right is to honestly account for the trade-offs between these two categories of innovation, recognizing three fundamental truths: society wants both, authors provide input that makes many of the relevant technologies more valuable, and technological advancement, in turn, typically makes copyrighted work more valuable too.

These issues are destined to repeat in a series of litigations over many years. Viacom's litigation against the video-sharing site YouTube is an example. The best reason to impose liability on YouTube is that it is in an enormously good position to filter for and in other ways discourage online infringement. The best reason to decline is that there will be some cost associated with filtering, and that cost might discourage future technologists from experimenting with similar products. My own view in that the case is comparable to my view here. To the extent that YouTube can discourage infringement at low cost—and it can[27]—copyright law could serve its many competing goals by requiring YouTube to take those steps. By contrast, where the costs of filtering would be crippling or where filtering would in other ways substantially interfere with legitimate amateur video distribution, copyright law could serve those same goals by acknowledging these harms and instead looking for other ways to reward and encourage authors.

Again, there is no formula for any of this, and a purely economic approach fails for lack of necessary data. There is still much to be gained, however, simply in recognizing copyright's role in not only encouraging the development and distribution of creative work but also in encouraging the development and deployment of related technologies. The more pervasive that recognition, the more courts and lawmakers will remember to address the relevant trade-offs, even if the efficient answer cannot be precisely quantified in a specific case or articulated in the form of a bright-line general legal rule.

Endnotes

1. For discussion, see Brief of Kenneth J. Arrow, Ian Ayres, Gary Becker, William M. Landes, Steven Levitt, Douglas Lichtman, Kevin Murphy, Randal Picker, Andrew Rosenfield, and Steven Shavell, as Amici Curiae in Support of Petitioners, *MGM Studios, Inc. v. Grokster Ltd.*, no. 04-480 (U.S. Sup. Ct., filed January 24, 2005).

2. See 17 U.S.C. §504(c), authorizing statutory damages, which is a type of financial consequence that might bear little relation to the efficient penalties the law ought to impose to deter bad acts.

3. See *Stewart v. Abend*, 495 U.S. 207, 236 (1990) and *Sony Corp. of America v. Universal City Studios, Inc.*, 464 U.S. 417, 448 (1984), which both describe fair use as an "equitable rule of reason."

4. *Michigan Document Services*, 1387, discusses how a permission-based copy shop system would work.

5. *Twentieth Century Music Corp. v. Aiken*, 422 U.S. 151, 156 (1975): "The immediate effect of our copyright law is to secure a fair return for an author's creative labor. But the ultimate aim is, by this incentive, to stimulate artistic creativity for the general public good" (citations omitted).

6. Transaction costs are well recognized as one of the central justifications for the fair use doctrine. The foundational paper is Gordon (1982).

7. See *Michigan Document Services*, 1387 (discussing permission fees), and *American Geophysical Union v. Texaco Inc.*, 802 F. Supp. 1, 7 (S.D.N.Y. 1992) (discussing the Copyright Clearance Center as an example of an intermediary that helps entities such as copy centers clear necessary permissions).

8. Interestingly, there was also a second constraint on licensing rates: the fact that, even after the decision in this case, individual students could make copies on their own and still invoke fair use. That is, the litigated case imposed liability on a formal copy center, in part on the intuition that copy centers operate on a sufficient scale that they could bear the costs of complying with the law and in complying would meaningfully alter author incentives. An individual student putting coins into a stand-alone copying machine, however, would not fall into the literal or intuitive scope of the opinion. This created a constraint on the total price charged by any copy center. If the copy center itself endeavored to mark up its product too much or if authors demanded too high a royalty, students would just opt for the less efficient loophole of copying for themselves.

9. My hesitation in this sentence comes only because it is easy to imagine the creation of a rights clearinghouse that would facilitate licensing of even these hard-to-license works. Indeed, enormous social value would be created were such a clearinghouse to be established because that clearinghouse could then facilitate all sorts of uses of these works above and beyond the index use that Google is here litigating. For now, however, such a clearinghouse does not exist. It would therefore be relevant to a court's analysis only if the court believed that, by denying fair use in this case, the court could increase the likelihood that such a clearinghouse would come into existence.

10. See H.R. Rep. no. 94-1476 (1976), 66; and *Harper & Row, Publishers, Inc. v. Nation Enterprises*, 471 U.S. 539, 549 (1985): "The statutory formulation of the defense of fair use in the Copyright Act of 1976 reflects the intent of Congress to codify the common-law doctrine."

11. See *Castle Rock Entertainment v. Carol Publishing Group, Inc.*, 150 F.3d 132, 141 (2d Cir. 1998): "The four listed statutory factors in §107 guide but do not control our fair use analysis and are to be explored, and the results weighed together, in light of the purposes of copyright" (citations and quotations omitted).

12. "[There is] no generally applicable definition [of fair use]," and "each case raising the question must be decided on its own facts" (*Harper & Row*, 560; quotations omitted). "The fair use test remains a totality inquiry, tailored to the particular facts of each case" (*Wright v. Warner Books, Inc.*, 953 F.2d 731, 740 (2d Cir. 1991).

13. See *Itar-Tass Russian News Agency v. Russian Kurier, Inc.*, 886 F. Supp. 1120, 1129 (S.D.N.Y. 1995): a "commercially exploitative use" is "one in which the copier directly and exclusively acquires conspicuous financial rewards from its use of the copyrighted material"; *A&M Records, Inc. v. Napster, Inc.*, 239 F.3d 1004, 1015 (9th Cir. 2001): even if the infringer does not actually profit from his bad act, "commercial use is demonstrated by a showing that repeated and exploitative unauthorized copies of copyrighted works were made [merely] to save the expense of purchasing authorized copies."

14. See *Salinger v. Random House, Inc.*, 650 F. Supp. 413, 425 (S.D.N.Y. 1986): "Publishers of educational textbooks are as profit-motivated as publishers of scandal-mongering tabloid newspapers. And a serious scholar should not be despised and denied the law's protection because he hopes to earn a living through his scholarship."

15. According to *Campbell*, a work is transformative when it does not "merely supersede the objects of the original creation."

16. See, e.g., *Rogers v. Koons*, 960 F.2d 301, 309 (2d Cir. 1992): "The first factor … asks whether the original was copied in good faith to benefit the public or primarily for the commercial interests of the infringer." One caveat to the above summary: some courts recognize a work as transformative only if the work is different from the original work in an expressive way. These courts do not accept evidence of just any new "purpose, meaning, or effect"; instead, they require a new expressive purpose, a new expressive meaning, or a new expressive effect. The rationale is that copyright law itself is designed to encourage expressive outputs and indeed itself refuses to protect valuable nonexpressive works such as databases and directories. See *Feist Publications, Inc. v. Rural Telephone Service Co.*, 499 U.S. 340 (1991). Some courts therefore think it appropriate to similarly distinguish expressive from nonexpressive work under the first fair use factor. Specifically, these courts refuse to recognize as transformative a work for which the new contribution is informational, organizational, or in some other way valuable but not expressive. See, e.g., *Infinity Broadcasting Corp. v. Kirkwood*, 150 F.3d 104 (2d Cir. 1998), which concluded that retransmission of radio broadcasts over telephone lines is not transformative despite the fact that the retransmission was for an entirely different, albeit nonexpressive, purpose.

17. Although, again, some courts will disagree, objecting that Google's use might be valuable, but it is not expressive. See, e.g., *Roy Export Co. v. Columbia Broadcasting System Inc.*, 503 F. Supp. 1137 (S.D.N.Y. 1980), *aff'd*, 672 F.2d 1095 (2d Cir. 1982): the broadcast of an unsponsored television program is still commercial because the broadcaster "stood to gain at least indirect commercial benefit from the ratings boost which it had reason to hope would … result" from airing the infringing program.

18. See *Perfect 10*, 720–21, and *Kelly v. Arriba Soft Corp.*, 336 F.3d 811, 818 (9th Cir. 2003), which both found that a search engine's infringement of copyrighted images was a transformative use despite the informational nature of the use.

19. See *Princeton University Press v. Michigan Document Services, Inc.*, 99 F.3d 1381, 1387 (6th Cir. 1996) (en banc), *cert. denied*, 520 U.S. 1156 (1997); *American Geophysical Union v. Texaco Inc.*, 60 F.3d 913, 930–31 (2d Cir. 1994), *cert. dismissed*, 516 U.S. 1005 (1995).

20. "The portions actually quoted were selected … as among the most powerful passages in those chapters" (566).

21. Thus, in *Campbell*, the Supreme Court very nearly excused under the fair use doctrine a musical parody that happened to be written as a rap. The court remanded to the lower court, however, for fear that the existence of a rap parody might significantly interfere with the original author's ability to license nonparodic rap versions in the future.

22. I must admit that this framing seems unduly narrow to me. The language seems to suggest that sufficiently abnormal or unlikely uses of copyrighted work should, as a rule, not be deemed infringement. The intuition, I suppose, is twofold: that these uses are themselves the sorts of creative accomplishments that copyright law should support and that unexpected uses do not matter anyway for author incentives. Both of those explanations fail, however. The first is too broad. Just because someone comes up with an unexpected use does not mean that the efficient result is to allow that person to capture all the value of that use rather than being required to share some of that value with the copyright holder whose work is incorporated. The second is just wrong. The incentive to create and disseminate creative work will indeed be dampened by a legal rule under which sufficiently new uses are given a free pass. Authors might not know what the relevant new uses will be, but they will expect that there will be such uses, and they will change behavior accordingly. I would therefore not adopt this particular language, but would instead consider the degree to which a use is unexpected as a relevant variable in the overall factor four analysis.

23. I say "in almost every case" rather than "in every case" because, in some cases, transaction costs would make it impossible for the accused infringer to pay even if the infringer wanted to.

24. See *American Geophysical Union*: "The vice of circular reasoning arises only if the availability of payment is conclusive against fair use" (931).

25. Discovery will reveal more information relevant to this discussion. For now, however, the already public contract between Google and the University of Michigan makes clear the mismatch between Google's incentives and author incentives. Google's contract imposes very few limitations on what Michigan does with the electronic copy of each book that Google provides to Michigan. Had the relevant copyright holders written the contract, surely they would have more carefully articulated Michigan's obligations to make sure that those electronic copies do not end up freely available on the Internet or in other ways abused.

26. It is possible to imagine that intermediaries would arise to search for new opt-out projects and opt out on behalf of participating copyright holders. That would reduce the overall costs of the opt-out approach, but it would be a complete waste from a social welfare perspective. In essence, the intermediaries would be recreating the opt-in approach currently in place, but doing so in a more cumbersome and costly manner.

27. For a fuller discussion of my view of the YouTube litigation, see Lichtman (2007).

References

Authors Guild. 2005. "Authors Guild Sues Google, Citing 'Massive Copyright Infringement.'" Press release (September 20), Authors Guild, New York.

Gordon, Wendy. 1982. "Fair Use as Market Failure: A Structural and Economic Analysis of the Betamax Case and Its Predecessors." *Columbia Law Review* 82:1600–1648.

Lichtman, Doug. 2007. "YouTube's Future—and Yes, It Has One." *Academic Advisory Council Bulletin* 2.3 (December), http://www.pff.org/issues-pubs/ip/bulletins/bulletin2.3youtubefuture.pdf.

Nimmer, Melville B. 2008. *Nimmer on Copyright*. New York: Lexis/Nexis.