The Google Book Settlement And The Fair Use Counterfactual

Matthew Sag, DePaul University
THE GOOGLE BOOK SETTLEMENT AND THE FAIR USE COUNTERFACTUAL

Matthew Sag*

ABSTRACT

This Article compares the pending settlement between Google and the representative author and publisher plaintiffs to the most likely outcome of the litigation the settlement resolves. This counterfactual provides a useful benchmark by which to assess the effects, and thus the merits, of the Google Book Search settlement. As well as providing a guide to understanding the settlement, this article identifies how the agreement mirrors the court’s most likely fair use ruling and how it deviates from it. Because the opt-out that fair use would have required has been replaced by the ability of copyright owners to opt-out of the class-action settlement, the Google Book debate has shifted away from the merits of book digitization, and refocused on questions of commoditization and control.

KEYWORDS

Google, Google Book Search, Google Settlement, Authors Guild, Antitrust, Copyright, Fair Use, Class Action.

* Associate Professor, DePaul University College of Law. The Final Fairness Hearing on the Google Book Settlement will take place on October 7, 2009. Revisions to this draft are likely following the hearing. Please address comments or corrections to msag@depaul.edu. My thanks to Jackson Cooper, Joshua McIntyre and Brian Murray for research. Thanks to James Grimmelmann, Peter Hirtle and Pamela Samuelson for their comments. Peter DiCola and I are working on a separate article in relation to the Google Book Settlement, he deserves substantial credit for his many contributions to this work. Responsibility for the content of this Article rests solely with the author. A video presentation paralleling various aspects of this article is available at http://www.youtube.com/watch?v=2FarIWX3oZI. Draft date August 20, 2009.
THE GOOGLE BOOK SETTLEMENT AND THE FAIR USE COUNTERFACTUAL

CONTENTS

Introduction ....................................................................................................3

Part I  The Counterfactual—What if There Were No Settlement? ..........9
  A. The Google Book Search Library Project .............................................9
  B. Copyright Infringement or Fair Use .................................................11
    1. The Purpose and Character of the Use ........................................13
    2. The Nature of the Copyrighted Work ...........................................18
    3. Amount and Substantiality .........................................................18
    4. Market Effect ..............................................................................20
    5. The Likely Fair Use Outcome ......................................................23

Part II  An Overview of the Google Book Settlement Agreement ..............24
  A. The Scope of the Settlement..........................................................25
  B. The Registry ..................................................................................27
  C. Financial Terms of the Settlement ..................................................31
  D. Uses and Permissions Under the Settlement ...................................32
    1. Book Search: Digitization & Display ........................................32
    2. Beyond Book Search ..................................................................35
    3. The Research Corpus ..................................................................40
    4. New Revenue Models ..................................................................41
  E. Book Classification ..........................................................................42
  F. Author-Publisher Procedures ............................................................45
    1. In-Print versus Out-of-Print .......................................................45
    2. Author-Controlled Books ............................................................47
    3. Who Controls the Exploitation of the Work? ...............................48
    5. Some Implications of the APP .....................................................54

Part III. Conclusion—Evaluating the Settlement ......................................55
THE GOOGLE BOOK SETTLEMENT AND THE FAIR USE COUNTERFACTUAL

INTRODUCTION

On October 28, 2008, Google announced that it had reached a settlement of the highly publicized class-action lawsuit brought by the Authors Guild and another equally important lawsuit brought by the American Association of Publishers. The U.S. District Court for the Southern District of New York will conduct a final fairness hearing for the class-action settlement beginning on October 7, 2009. If the settlement is ultimately approved by the court, it will end four years of uncertainty and clear the way for Google to continue to digitize millions of copyrighted library books and further develop the Google Book Search Library Project. The question remains, however, should the settlement should be approved?


2 Order dated April 28, 2009, Author’s Guild et al. v. Google, Inc., Civil Action No. 05 Civ. 8136 (DC).

The Google Book Search Settlement Agreement (the “Settlement”) is a complicated document with far reaching implications. It is also subject to many different frames of analysis. Antitrust experts will no doubt be asking whether the market for online book searching will become more or less competitive in light of the Settlement and how the broader Internet search market will be affected by the agreement. Class-action lawyers may question the Settlement should be approved under Rule 23(e) and whether it constitutes an appropriate use of the judicial process—i.e. whether it is both procedurally fair and substantively fair, adequate and reasonable. Librarians will speculate on what the Settlement portends for the future of their institutions and indeed for the future of the book as a physical medium for the communication of knowledge. From a social welfare perspective, the immediate question is simply whether society as a whole is better off with the Settlement than without it.

How then should we evaluate the proposed Settlement? We could compare the Google Book Search project that will result from the Settlement to some free equivalent established by the library of Congress or a similar government agency. However, mass digitization has clearly not been a priority for the Library of Congress in the past and it is unclear to what extent Google’s own efforts at book digitization could actually impede subsequent digitization efforts by government or non-government actors. Alternatively, we could compare the Settlement negotiated by the parties to a legislative solution that clarified rights and liabilities with respect to book digitization or perhaps created some kind of safe harbor. This would no doubt be a useful vein of enquiry, however without a definitive legislative proposal it is difficult to make this comparison in the abstract and one risks comparing the actual Settlement to an idealized, well-written, unbiased, capture-free legislative solution.

This Article takes a different approach, comparing the Settlement to the most likely outcome of the litigation it resolves. This counterfactual is

---


5 The district court has considerable discretion in determining whether settlement is fair and reasonable. Bryan v Pittsburgh Plate Glass Co. (PPG Industries, Inc.), 494 F2d 799 (3d Cir. 1974).

6 Robert Darnton, supra note 3.
worth exploring because it allows us to concentrate on the merits of the Settlement per se, rather than the merits of the Google Book Search project more broadly. The fair use issues in relation to the Google Book Search Library Project have been largely misunderstood. Google was never likely to receive the courts unqualified approval for its massive digitization effort. Instead, the most likely outcome of the litigation was that book digitization would qualify as a fair use so long as copyright owners were given the opportunity to opt-out of inclusion in the book database and that opportunity was made freely available at a cost that was essentially trivial.

Assuming that any fair use ruling would have been conditioned on the provision of an opt-out, many of the terms of the Settlement do not differ significantly from the most likely outcome of the litigation. Essentially, the opt-out that fair use would likely have required has been replaced in the Settlement by the ability of copyright owners to opt-out of the class-action and the significant opt-out and modification opportunities within the Settlement itself. The aspects of the Settlement which allow Google to continue to operate its book search engine substantially in its current form should not be controversial; they essentially mirror the court’s most likely fair use ruling if the case had gone to trial. The initial digitization of books, the processing and analysis of metadata and basic reporting functions of the book search engine were likely to constitute fair use and these are also allowed under the terms of the Settlement.

It is not surprising that, in the wake of the Settlement, the debate has shifted from the merits of digitization to commoditization and control. These are the issues where the Settlement differs sharply from the predicted fair use ruling.

First, the Settlement permits Google to engage in a significant range of uses beyond fair use. In addition to allowing book previews the Settlement also envisages new services including a “Consumer Purchase” model, “Institutional Subscriptions” and various other “New Revenue Opportunities” that remain subject to negotiation between Google and the rightsholders. These new features take the database well beyond the realm of mere indexing and identification which enhances the value of paper collections. Electronic book purchases, corporate and educational subscriptions and various other “New Revenue Opportunities” will allow

---

7 See infra, Part I.
8 See infra, Part I-1 and Part I-4.
9 See infra, Part II.
for the complete electronic distribution of books that substitutes entirely for paper-based mediums.

Second, the Settlement provides for initial cash payments by Google to the copyright owners and a fairly generous revenue sharing agreement, neither of which would have been required under a fair use ruling. These payments can be viewed as either appropriate compensation for book digitization, a *quid pro quo* for the expanded uses provided for under the agreement, or both. This may not be the only *quid pro quo* brought about by the Settlement. One troubling aspect of the deal between Google and Authors Guild and the Association of American Publishers is that it greatly enhances the power of these organizations. Before the Settlement, the Authors Guild spoke for a mere 8000 authors, if the agreement is approved that same organization, as the representative of the author subclass will effectively speak for millions.

Third, the agreement creates a new set of institutional arrangements that will govern the relationship between Google and the copyright owners covered by the Settlement. The foundations of this new institutional framework are the Settlement agreement itself, the creation of a collective rights management organization called the “Book Rights Registry” and the “Author-Publisher Procedures”.

Compared to a fair use ruling, the Settlement gives copyright owners a greater degree of flexibility to manage the types of uses that are made of their works and the degree to which their works are made available for preview and display. While in theory copyright owners would still have been able to contract around a fair use ruling, their ability to do so easily and effectively is greatly enhanced by the machinery put in place by the Settlement. By establishing clear mechanisms and default rules, the agreement effectively lowers transaction costs for individual copyright owners in a number of ways. More importantly, the creation of the Registry allows rightsholders to exploit their works collectively even more efficiently. The Registry will function as a collective rights institution that will facilitate rightsholder action and also stand in the shoes of inactive rightsholders.

---

10 See *infra* Part II-C.
12 See *infra* Part II-D.
13 *Id.*
Another key institutional feature of the settlement is the Author-Publisher Procedures which play a substantial role in determining who controls the exploitation of a work within the Google Book universe and also who benefits from that exploitation.\textsuperscript{14} These procedures enhance the coordinating function of the Settlement by streamlining the incorporation of existing author-publisher contractual terms into the framework of the Google Book universe. However, where an existing author-publisher contract gives both parties some control over electronic exploitation, or simply fails to make any provision for electronic rights, the Author-Publisher Procedures effectively overwrite those contracts. These new terms do not systematically disadvantage either authors or publishers, but there may well be individual authors or publishers who would strenuously object to them.

The fourth area in which the Settlement differs from the likely fair use outcome relates to the accessibility, commoditization and control of orphan works. The Settlement gives Google, and the public, significantly greater access to orphan works. Whereas there is no strong fair use argument for allowing substantial previews of up to 20 percent of a book’s contents, the Settlement makes this “Standard Preview” the default for most types of books.\textsuperscript{15} This default setting ensures a much greater level of access to orphan works than could have been possible merely relying on fair use and other defenses to copyright infringement. More controversially, orphan works will be commoditized as Consumer Purchases and Institutional Subscriptions and the profits from this commoditization will flow in part to the Registry, and to registered authors and publishers.

The issue of orphan works is complicated by the ambiguity and elasticity of the term itself. Orphan works are works for which the copyright owner cannot be readily identified and located by someone who might need permission to use the work.\textsuperscript{16} The ambiguity here pertains to how much effort a user should be expected to exert to track down wayward parents and negotiate with them. Some copyrighted works will never be identifiable with reasonable certainty either because initial authorship is ambiguous or because the copyright ownership has been passed along a chain of title to a complex too unravel. In contrast to these “pure orphans” there are “contextual orphans”, works for which the copyright owner could be found, but for which the expenses of so doing is prohibitive or unjustifiable on a cost benefit analysis. Potential users of both pure and contextual orphan

\textsuperscript{14} See \textit{infra}, Part II-F.
\textsuperscript{15} § 4.3(b)(i)(1) (Standard Preview), see \textit{infra} Part II-D(1) for further discussion.
\textsuperscript{16} Copyright Office Report on Orphan Works.
works often choose not to use or display such works out of concern that they may be subject to injunctions and statutory damages if the owner of an orphan work subsequently comes into the picture.

The orphan works problem is particularly acute for library and archive digitization projects and the problem has been exacerbated in recent years due to both supply-side and demand-side considerations. On the supply-side, the extension of the term of copyright protection, the elimination of formalities such as notice and registration, and the retroactive restoration of some foreign works previously in the public domain, have all increased the domain of copyright protection at the expense of the public domain. At the same time, digitization has increased the demand for access to more obscure works. Tracing the ownership of copyright requires more than just tracking down the original author—copyrights are transferred by the operation of certain statutory provisions, private contracts, corporate mergers, divestitures and bankruptcies, and as part of deceased estates. As the Library Copyright Alliance notes:

[a] person seeking permission to use an older work needs to untangle the complicated history of mergers and acquisitions of the original publisher to identify the corporation that now might own the rights to the work. This corporation, however, might not have adequate records concerning all the works to which it holds title. And since the person typically is seeking to use the work for at most nominal compensation, the corporation has no incentive to invest adequate resources in locating the records concerning the work.

Google’s initial approach the difficulty of proactively clearing rights with millions of different authors and their descendants was to rely on the fair use doctrine and the ability of copyright owners to opt-out as a substitute for express permission. This was not only a high-risk legal strategy; it also significantly limited the types of uses Google could make of orphan works. The Settlement, in contrast, gives Google an express license from the owners of orphan works, albeit a license that remains subject to substantial opportunities for active copyright owners to modify key terms or opt-out of key uses altogether.

18 Id.
One of the most pronounced criticisms of the Settlement is that it will create an “orphan works monopoly.”19 The core of this allegation is that while a fair use ruling in Google’s favor would have unblocked access to orphan works for a broad class of users, the Settlement only solves the orphan works problem for Google.20 Although it is true that the Settlement creates no fair use precedent, the distinct set of legal entitlements which will vest on Google if the Settlement is approved will only translate to a monopoly if they establish a barrier to entry in the book search market. The Settlement may not constitute a barrier to entry because in some respects Google’s competitors will have a considerable second mover advantage. Although it only Google can fully commoditize orphan works, the Registry, created at Google’s expense, provides Google’s competitors with an easier route to reaching an agreement with a large and diverse group of active copyright owners. Moreover the Settlement defines an effective benchmark of reasonable terms and conditions that would-be competitors may seek to emulate.21

This Article is organized as follows. Part I establishes the counterfactual basis for comparison and assesses the extent to which Google’s assertion of fair use would have trumped the copyright owners’ claim of infringement. Part II reviews the essential features of the Settlement Agreement in considerable detail. By way of conclusion, Part III then compares the Settlement to the likely outcome under the fair use doctrine. Addressing these three issues does not, by itself, answer the ultimate question of whether the Settlement should be approved. But hopefully it brings us several steps closer to that objective.

PART I THE COUNTERFACTUAL—WHAT IF THERE WERE NO SETTLEMENT?

A. The Google Book Search Library Project

The Google Book Search Library Project (“GBS”) is an attempt to build a comprehensive search engine that will allow full text searching inside millions of books.22 Google has been obtaining the raw data for this search

20 Picker, Id.
21 The counterfactual analysis undertaken here is an essential predicate to resolving these antitrust questions, although that task itself is beyond the scope of this Article.
engine in electronic form from a variety of book publishers since 2004.\textsuperscript{23} However, because most books published over the last few hundred years are not available in electronic form, Google has also been working with the various libraries to digitize their paper collections shelf by shelf, page by page—a painstaking, time-consuming and expensive process.\textsuperscript{24}

Once a book is digitized, Google can treat it the same way it treats an HTML webpage. Google stores the digitized volumes on its computers and performs intense data analysis in order to generate what is essentially a vastly improved library card catalogue system for the digital age. Once digitized and analyzed, a book is then available for reporting in response to user search queries. In its current form, Google book search reports either limited bibliographic information about a book, small snippets of text from a book, a limited preview which displays some substantial sections of a book, or, if a book is not subject to copyright in United States, it makes the entire book available for digital download.\textsuperscript{25}

Standing alone, the standard way that Google reports of search results is unlikely to amount to copyright infringement for two principle reasons. First, the basic bibliographic information that Google provides such as title, author and subject matter, are all clearly uncopyrightable facts. Through the idea-expression distinction, copyright law protects the expressive elements of the author’s work while guaranteeing subsequent authors the necessary breathing space to make their own contributions by adding to, re-using, or re-interpreting the facts and ideas embodied in the original work.\textsuperscript{26} Copyright law clearly distinguishes between facts and their expression and it is well settled that data is not subject to copyright protection: one can extract and reproduce facts, names and dates from a newspaper article, or ideas and processes from an instructional text, without infringing the author’s copyright.\textsuperscript{27}

\textsuperscript{24} \textit{Id.} Note that this Article does not address whether Google’s practice of making digitized versions of books available to participating libraries could itself amount to copyright infringement.
\textsuperscript{25} Google, \textit{What you'll see when you search on Google Book Search}, available at \url{http://print.google.com/googleprint/screenshots.html}.
\textsuperscript{26} 17 U.S.C. § 102(b) provides: “In no case does copyright protection for an original work of authorship extend to any idea, procedure, process, system, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or embodied in such a work.”
Second, the very brief snippets or quotations that Google displays are in all likelihood, too brief, too fragmented and too insubstantial to constitute a reproduction of an entire copyrighted book. Although the Copyright Act prohibits certain forms of reproduction, the actual threshold of infringing reproduction is left undefined by the Act.\footnote{17 U.S.C. § 106 (setting out the exclusive rights of the copyright owner).} Thus the question of how much copying is too much turns on common law tests of what is usually referred to as “substantial similarity.”\footnote{Meaning protectable similarity, not merely the similarity that may prove that work B was actually copied from work A. See, Laureyssens v. Idea Group, Inc., 964 F.2d 131, 140 (2d Cir. 1992); Tufenkian Import/Export Ventures, Inc. v. Einstein Moomjy, Inc., 338 F.3d 127, 131 (2d Cir. 2003).} Where a fragment of the original has been copied exactly, courts determine whether the copying amounts to infringement “by considering the qualitative and quantitative significance of the copied portion in relation to the plaintiff’s work as a whole.”\footnote{Newton v. Diamond, 388 F.3d 1189, 1195 (9th Cir. 2004) (citing Worth v. Selchow & Righter Co., 827 F.2d 569, 570 n. 1 (9th Cir. 1987); Jarvis v. A&M Records, 827 F. Supp. 282, 289-90 (D.N.J. 1993); 4 Nimmer § 13.03[A][2], at 13-47 to 48).} The snippets of text Google displays in response to a search query extremely unlikely to cross over this threshold of substantiality. Google Book also provides extended preview displays of various books; however, these are all subject to express licenses from the relevant copyright owners, usually in through one of Google’s Publisher Partner Agreements.\footnote{Google, \url{https://books.google.com/partner/}} Google’s book search engine also offers complete access to works known to be in the public domain.\footnote{Google, \url{what you'll see when you search on Google Book Search}, available at \url{http://print.google.com/googleprint/screenshots.html}.}

**B. Copyright Infringement or Fair Use**

Lawyers for the Authors Guild and the plaintiff publishers have argued that even if the search results which Google reports to users do not violate the Copyright act—a point they do not concede—scanning millions of copyrighted library books in order to build that search engine in the first place certainly does.\footnote{See, Author’s Guild et al. v. Google, Inc., First Amended Class Action Compliant, paragraphs 3 and 4.} The copyright owner has an exclusive right to reproduce his or her work in copies under the Copyright Act.\footnote{17 U.S.C. § 106(1)(the exclusive right to reproduce the copyrighted work in copies).} There is no doubt that digitally scanning a book and storing that scan on a computer hard drive where it can be retrieved and reprinted amounts to making a copy
in the ordinary sense of the word. Google’s primary counterargument to this charge of infringement is that its new technology is protected by the fair use doctrine.

The fair use doctrine plays a vital role balancing “the interests of authors and inventors in the control and exploitation of their writings and discoveries on the one hand, and society’s competing interest in the free flow of ideas, information, and commerce on the other hand” Fair use is a flexible standard which limits the scope of copyright protection and renders certain actions relating to copyrighted works non-infringing. Accordingly, incidental copying in the service of quotation, criticism, illustration, comment, parody and simple clarification have long been recognized as fair use.

The balancing function of the fair use doctrine is critically important where the copyright owner withholds permission for reasons that we as a society find unacceptable, such as to stifle parody, criticism, or social debate. However, the doctrine is not limited to the suppression of dissenting opinions. Fair use is necessary even when copyright owners are purely commercially motivated because licensing and other private ordering mechanisms do not provide a solution for high exchange costs, high information costs and strategic behavior. The fair use doctrine is particularly important in situations where the costs of obtaining permission outweigh the benefits of the use.

Judicial application of the fair use doctrine is often regarded as highly fact-specific and unpredictable. Nevertheless, courts predictably frame their fair use decisions in terms of four non-exhaustive statutory factors. These factors are (1) the purpose and character of the defendant’s use; (2) the

35 “Copies” are defined under the Act as “material objects … in which a work is fixed… and from which the work can be perceived, reproduced, or otherwise communicated.” 17 U.S.C. § 101.
41 See Dellar v. Samuel Goldwyn, Inc., 104 F.2d 661 (2d Cir. 1939) (describing fair use as one of the most troublesome doctrines in the whole law of copyright). This unpredictability may be overstated. See, Pamela Samuelson, Unbundling Fair Uses, 77 FORDHAM L REV 2537 (2009).
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.\footnote{17 U.S.C. § 107. In reality, the Section 107 factors are neither complete, nor are they individually or cumulatively determinative. Even the notion that there are four factors is misleading: beneath the statutory factors lies an amalgamation of interconnected meta-factors, sub-factors and presumptions. Bond v. Blum, 317 F.3d 385, 394 (4th Cir. 2003) ("These factors are not meant to be exclusive, but rather illustrative, representing only general guidance about the sorts of copying that courts and Congress most commonly have found to be fair uses.") (citations omitted); Universal City Studios, Inc. v. Sony Corp. of America, 480 F. Supp. 429, 448 (C.D. Cal. 1979) ("The factors are illustrative, not definitive.")}{17 U.S.C. § 107(1).

1. The Purpose and Character of the Use

The first fair use factor is “the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes.”\footnote{43} The case law under the first factor is dominated by two considerations, whether the use is transformative and whether the use is commercial.\footnote{44} According to the Supreme Court’s most recent fair use decision, \textit{Campbell v. Acuff-Rose}, the first factor turns primarily on “whether and to what extent the new work is ‘transformative’”\footnote{45} A transformative use is one that, compared to the original creation, “adds something new, with a further purpose or different character.”\footnote{46} Transformative uses are protected by the fair use doctrine because they lead to the creation of new works which do not substitute for the author’s original expression.\footnote{47} For example, parodies are considered to be transformative because they shed light on an existing work while at the same time creating something new.\footnote{48} A claim to transformative use is most obvious when the work itself is literally transformed; however, in many cases courts have held that other uses that do not literally change the work are nonetheless sufficiently different to constitute fair use.

Google can point to significant precedents in relation to similar technology where the courts have found intermediate copying constitutes fair use. In a recent decision about plagiarism detection software, the Fourth Circuit held

\footnote{17 U.S.C. § 107(1).}{43}


\footnote{\textit{Id.}}\textsuperscript{46}

\footnote{\textit{Id.} at 579.}{47}

\footnote{\textit{Id.}}\textsuperscript{48}
that copying and archiving student term papers was transformative because the papers were used “for an entirely different purpose, namely, to prevent plagiarism and protect the students’ written works from plagiarism.” In several reverse engineering cases, courts have held that copying computer software in order to discover functional compatibility requirements is a fair use. Likewise, in the context of visual search engines, the Ninth Circuit has also held that using thumbnail versions of copyrighted photographs as part of a menu of search results is transformative and constitutes a fair use. In Perfect 10. v. Amazon.com, the Court of Appeals concluded that the use of these thumbnail photos may even be “more transformative than a parody because a search engine provides an entirely new use for the original work, while a parody typically has the same entertainment purpose as the original work.”

There are good reasons to think that this view of transformative use would prevail in the Second Circuit where the Google Book case is being litigated. A number of significant cases from that circuit have declined to find a transformative use when the defendant has done no more than locate a new way to exploit the creative virtues of the original work. However, the Second Circuit has recognized the need to extend fair use beyond artistic and critical transformations to situations where the defendant uses a copyrighted work in a different context to serve a fundamentally different function than the original. Moreover, although the most relevant search

---

49 A.V. v. iParadigms, LLC, 562 F.3d 630, 640 (4th Cir. 2009).
50 Sega Enterprises Ltd. v. Accolade, Inc., 977 F.2d 1510 (9th Cir. 1993); Sony Computer Entertainment, Inc. v. Connectix Corp., 203 F.3d 596, 607 (9th Cir. 2000). In Sega Enterprises Ltd. v. Accolade, Inc., the Court of Appeals for the Ninth Circuit held that “Accolade copied Sega’s code for a legitimate, essentially non-exploitative purpose” made its use essentially fair under the first factor. Sega 977 F.2d 1522-23. In Sony Computer Entertainment, Inc. v. Connectix Corp., a case with similar facts, the court concluded that the defendant’s Virtual Game Station was transformative because it created a new gaming platform and was itself “a wholly new product”. Sony Computer Entertainment, Inc. v. Connectix Corp., 203 F.3d 596, 606-607.
51 Kelly v. Arriba Soft Corp., 336 F.3d 811, 821 (9th Cir. 2003); Perfect 10, Inc. v. Amazon.com, Inc., 487 F.3d 701, 721 (9th Cir. 2007)
52 See e.g., Davis v. The Gap, 246 F.3d at 174 (use of plaintiff’s eyewear in a clothing advertisement not transformative because it was “worn as eye jewelry in the manner it was made to be worn”); Castle Rock Entm’t, 150 F.3d at 142-43 (Seinfeld quiz book not transformative when its purpose was to repackagel the television show Seinfeld to entertain Seinfeld viewers); Ringgold v. Black Entm’t Television, Inc. 126 F.3d 70, 79 (2d Cir. 1997) (copy of plaintiff’s painting used as decoration for a television program’s set not transformative because it was used for “the same decorative purpose” as the original).
53 Bill Graham Archives v. Dorling Kindersley Ltd., 448 F.3d 605, 609-610 (2d Cir. 2006) (finding transformative use when defendant’s purpose in using copyrighted concert poster was “plainly different from the original purpose for which they were created” and “a
engine precedents do not hale from the Second Circuit, both Amazon and Kelly have been cited by U.S. district courts in the Second Circuit with apparent agreement on this very point.\textsuperscript{54}

If the Google book search engine is transformative, does it matter that it is also commercial? The short answer is “no”. In its 1994 decision of \textit{Campbell v. Acuff-Rose}, the United States Supreme Court clearly rejected the notion that commerciality by itself had any “hard presumptive significance.”\textsuperscript{55} Instead, the Court adopted a sliding scale to commercial use, arguing that because “the goal of copyright, to promote science and the arts, is generally furthered by the creation of transformative works… the more transformative the new work, the less will be the significance of other factors, like commercialism, that may weigh against a finding of fair use.”\textsuperscript{56}

There are countless examples of commercial fair use. There are also a substantial number of cases where commercial use by an Internet search engine was deemed to be fair use in analogous circumstances. In \textit{Kelly v. Arriba Soft Corp.}, an image search case preceding \textit{Perfect 10}, the Ninth Circuit ruled that the replication of copyrighted images in thumbnails would not substitute for the full-sized images.\textsuperscript{57} The court in \textit{Perfect 10} likewise concluded that Google’s thumbnail-representations were unlikely to interfere with the market for Perfect 10’s original expression.\textsuperscript{58} The court expressly rejected the application of any commerciality inference or presumption noting that “this presumption does not arise when a work is transformative because market substitution is at least less certain, and market harm may not be so readily inferred.”\textsuperscript{59} In sum, the commercial nature of Google’s service proves to be little more than a red herring in relation to the first factor.

\begin{footnotesize}
\begin{itemize}
\item\textsuperscript{55} Id. at 585.
\item\textsuperscript{56} Id. at 579. As Barton Beebe notes, while commentators have assumed that the commerciality presumption was finally discarded in \textit{Campbell}, it remains a tenacious meme in the court of public opinion. \textit{See}, Barton Beebe, \textit{An Empirical Study of U.S. Copyright Fair Use Opinions, 1978-2005}, 156 U. Pa. L. Rev. 549, 598 (2008).
\item\textsuperscript{57} Kelly v. Arriba Soft Corp., 336 F.3d 811, 815 (9th Cir. 2003).
\item\textsuperscript{58} Perfect 10, Inc. v. Amazon.com, Inc., 487 F.3d 701, 724 (9th Cir. 2007).
\item\textsuperscript{59} Id. (citing \textit{Campbell v. Acuff-Rose Music}, 510 U.S. at 591).
\end{itemize}
\end{footnotesize}
Google’s decision to give authors the ability to opt-out of its book search engine should also have a significant impact on the “purpose and character of the use” in any fair use analysis. This opt-out flips the usual default rule from exclusion to the inclusion. In essence, it allows Google to continue scanning books and incorporating them into the book search engine until it receives an instruction to the contrary from the copyright owner. Without such an opt-out mechanism, clearing rights for Google’s book search engine would prove to be a very expensive proposition indeed.

How expensive? There are an estimated 18 million books in the combined collections of the libraries participating in the Google Book project. Approximately 10.5 million of these books are unique—i.e., they are only held by one of the participating libraries. It is estimated that slightly less than 20% of these works were published before 1923 and thus present no copyright issues. That leaves about 8.4 million books with some potential copyright constraint. Even if the average clearance cost (the cost of determining the status of the book, finding the relevant copyright owners and negotiating a license) was a mere $200, the total cost of rights clearance before any royalties have been paid would be $1.68 billion.

This explains why the opt-out is beneficial for Google, but how exactly does the opt-out make Google’s fair use case stronger? The first thing to understand here is that opt-outs are not that unusual. They are in fact common in a number of scenarios, and particularly characteristic of copy-reliant technologies. Traditional Internet search engines must navigate and assimilate information from literally billions of webpages. In theory, the software programs that crawl across the Internet could query every page for permission to enter, permission to copy, and permission to include in a search engine. If this were the case, a large number of web pages would probably post the appropriate invitations in their HTML code, however, a large number would probably fail to do so through ignorance or indifference. In actuality, the Internet default rule of openness prevails, and the burden of opting out of search engine inclusion, indeed of accessibility in general, is placed upon those with minority preferences. This state of

---

61 Id.
63 Id. (Identifying and discussing the significance of the emergence of copy-reliant technology).
affairs is generally accepted because those who wish to opt-out are not substantially inconvenienced. Indeed, it only takes the a few lines of code to activate the “robots exclusion protocol” which is the Internet equivalent of a “no trespassing” sign.\textsuperscript{64} The robots exclusion protocol allows the website controller to exclude search engine robots altogether, or to specifically tailor their permissions to suit their needs.\textsuperscript{65}

Why should the choice between opt-in and opt-out make any difference at all? In a Coasian world (one with no transaction costs) the choice between an opt-in and opt-out default is entirely inconsequential.\textsuperscript{66} As every student of the \textit{Coase Theorem} knows, in a world without transaction costs, the specific allocation of costs and benefits is unimportant because individuals will bargain to an efficient outcome regardless of their initial entitlement.\textsuperscript{67} Accordingly, if three out of four authors prefer to be included in the Google book search engine and one in four prefers not to be, then either 75\% of authors will opt-in, or 25\% of authors will opt-out.

However, as every student of the \textit{Coase Theorem} should also know, the absence of transactions costs is merely a thought experiment—in actuality, reallocation and enforcement are often very costly.\textsuperscript{68} Thus, in the real world, opt-out mechanisms such as the robots exclusion protocol can be extremely important in resolving large transaction costs problems. Opt-out mechanisms should not be seen as displacing private ordering; on the contrary, they are the means of private ordering. When transaction costs are otherwise high, low cost opt-out mechanisms can play a critical role in preserving a default rule of open access while still allowing individuals to have their preferences respected.\textsuperscript{69}

In the context of the Google Book project, it seems likely that the opt-out makes the “purpose and character” of the use more fair, subject to three conditions: (i) a majority of copyright owners either favor or do not object to inclusion in the search engine; (ii) the expected cost of proactively obtaining permissions is very high; and (iii) the costs of taking advantage of the opt-out are sufficiently small such that the autonomy of the minority is preserved. It cannot be said with certainty that Google would have been

\begin{itemize}
  \item \textsuperscript{64} Id.
  \item \textsuperscript{65} Id.
  \item \textsuperscript{67} Id.
  \item \textsuperscript{68} See Ronald H. Coase, \textit{Essays on Economics and Economists} 11 (1994). (“I tend to regard the Coase Theorem as a stepping stone on the way to an analysis of an economy with positive transaction costs.”)
  \item \textsuperscript{69} Sag, \textit{supra} note 62.
\end{itemize}
successful in establishing these three elements, however it seems more likely than not.

2. The Nature of the Copyrighted Work

On its face, the second fair use factor, “the nature of the copyrighted work” suggests that “some works are closer to the core of intended copyright protection than others, with the consequence that fair use is more difficult to establish when the former works are copied.” In general, and the allowable scope of fair use is thought to be greater with respect to published works and factual or informational works and narrower with regard to unpublished works and more expressive and creative works. The first of these distinctions appears to favor the Google, as the overwhelming majority of library books are published. As to the second distinction, it is obvious that the books being scanned by Google will occupy every conceivable point on the continuum between informational and expressive. Thus, this factor could tilt in Google’s favor in some cases, while being neutral or adverts in others.

Nonetheless, the second factor was never likely to be particularly instructive for a court considering Google’s assertion of fair use. The Supreme Court has noted that the second factor is rarely “likely to help much in separating the fair use sheep from the infringing goats in a parody case”. Subsequent decisions have confirmed that the utility of the second factor is limited where the creative work is being used for a transformative purpose. Like so many other cases, the significance of the second factor would probably have been overwhelmed by other considerations.

3. Amount and Substantiality

The third factor that a court must consider in any fair use determination is “the amount and substantiality of the portion used [by the defendant] in relation to the copyrighted work as a whole.” The question of ‘how much copying is too much for fair use to apply’ is beguilingly simple. In truth, the third factor cannot be reduced to any simple mechanical quantification. There is simply no magic number in fair use cases because the extent of permissible copying varies according to the purpose of the copying and the

---

71 NXIVM Corp. v. Ross Inst., 364 F.3d 471, 480 (the parties did not dispute that because the copyrighted work was unpublished, the second fair-use factor favored the plaintiffs).
73 Bill Graham Archives v. Dorling Kindersley Ltd., 448 F.3d 605, 612. See also, Blanch v. Koons, 467 F.3d 244, 257 (2d Cir. 2006).
effect of the copying on the rightsholder. The key enquiry under the third factor is not simply how much of the copyright owner’s work was used by the defendant, but rather, what proportion of the work’s expressive value had been appropriated. Transformative uses in general, and non-expressive uses in particular, do not tend to substitute for the value of the author’s original expression and are thus regarded favorably under the third factor.

Google can point to a significant body of precedent where total literal copying has been accepted as fair use. For example, in the seminal fair use case of Sony v. Universal the Supreme Court held that consumer time shifting of broadcast television constituted a fair use in spite of the fact that it clearly required complete as opposed to partial copying. Likewise, the Ninth Circuit has suggested that converting music files from a compact disc format to MP3 files also constitutes a fair use. In the recent case of Bill Graham Archives v. Dorling Kindersley the Second Circuit found that the complete reproduction of various Grateful Dead concert posters on a reduced scale as part of an illustrated Grateful Dead biography was protected by the fair use doctrine. Reverse engineering cases such as Sony v. Connectix and Sega v. Accolade, and Internet search engine cases such as Kelly v. Arriba Soft, Perfect 10 v. Amazon.com and Field v. Google are even more favorable for Google. In each of these cases, the relevant court determined that the total literal copying required by the search engine technology at issue was justified under the fair use doctrine. These cases appear to be exactly on point with respect to Google’s book search engine.

78 Recording Indus. Ass’n of Am. v. Diamond Multimedia Sys., 180 F.3d 1072 (9th Cir. 1999). See also, Lewis Galoob Toys, Inc. v. Nintendo of Am., Inc., 964 F.2d 965, 971 (9th Cir. 1992).
79 Bill Graham Archives v. Dorling Kindersley Ltd., 448 F.3d 605, 613 (2d Cir. 2006) (holding that total copying “does not necessarily weigh against fair use because copying the entirety of a work is sometimes necessary to make a fair use of the image”). See also Mattel Inc. v. Walking Mt. Prods., 353 F.3d 792, 803 n.8 (9th Cir. 2003) (holding that “entire verbatim reproductions are justifiable where the purpose of the work differs from the original”), Nunez v. Caribbean Int’l News Corp., 235 F.3d 18, 24 (1st Cir. 2000) (concluding that to copy any less than the entire image would have made the picture useless to the story);
4. Market Effect

The fourth fair use factor is “the effect of the use upon the potential market for or value of the copyrighted work.” To determine whether Google’s massive book scanning enterprise is likely to put a dent in the market for or value of the works being digitized necessitates defining a\ priori precisely what the market constitutes. If this case had gone to trial on the issue of copyright infringement, Google would have no doubt argued that its library project, far from harming copyright owners, would benefit them enormously by stimulating demand for their books. The author and publisher plaintiffs may not dispute this, but they would almost certainly argue that the digital library Google is attempting to create constitutes a viable new market for their works.

How would a court determine if the scanning required to stoke the fires of Google’s book search engine is part of the copyright market belonging to authors and publishers? Copyright owners have often claimed that new opportunities created by technological change are part of the intrinsic value of, or potential future market for, their works. Sometimes these assertions are successful—e.g., A&M Records, Inc. v. Napster, Inc., —and sometimes they are not—e.g. Sony Corporation of America v. Universal City Studios, Inc., Recording Industry Association of America v. Diamond Multimedia Systems.

As a conceptual point, it is self-evident that the “potential market” referred to in the statute cannot be taken to mean every possible use of the copyrighted work. If every use is part of the copyright owner’s market, then every claimed fair use has a potential market effect and the enquiry becomes meaningless. The argument simply proves too much—the law continues to draw a distinction between that which might be licensed and

---

82 A&M Records v. Napster, Inc., 239 F.3d 1004 (9th Cir. 2001); Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417 (1984); Recording Indus. Ass'n of Am. v. Diamond Multimedia Sys., 180 F.3d 1072 (9th Cir. 1999).
83 American Geophysical Union v. Texaco Inc., 60 F.3d 913, 930, n.17 (2d Cir. 1994) (“were a court automatically to conclude in every case that potential licensing revenues were impermissibly impaired simply because the secondary user did not pay a fee for the right to engage in the use, the fourth fair use factor would always favor the copyright holder.”) (citations omitted). See also, Leval, supra note 45, at 1124 ("[b]y definition every fair use involves some loss of royalty revenue because the secondary user has not paid royalties").
that which must be licensed.\textsuperscript{84} A copyright owner cannot prove a market effect by demonstrating that if it wins the lawsuit, the defendant, and people like the defendant, will have to pay for permission. Likewise, a defendant cannot prove that there is no market effect by demonstrating that if it wins the lawsuit it will be under no obligation to pay the copyright owner, and thus there will be no market for the particular use at issue. This kind of circular logic proves nothing for either party. The fact is, to determine whether there is a market effect, a court must determine whether the use in question is something that falls within the legitimate sphere of copyright owner control or not.\textsuperscript{85} Courts will look to previous fair use cases to make this determination.

Although the fair use doctrine is frequently derided as unpredictable, there are actually a number of clear principles which emerge from the case law which are relevant to this analysis. It is clear that the rights of the copyright owner are not strictly limited by past exploitation.\textsuperscript{86} In other words, just because the copyright owner has not ventured into a particular market in the past, does not necessarily limit them from doing so in the future. It is also well established that speculative assertions of impending harm are insufficient.\textsuperscript{87} Perhaps most importantly, there is now a rich body of case law illustrating that the kinds of market effects courts will recognize are limited in several important ways.

The first limiting principle is simply likelihood: “[t]he market for potential derivative uses includes only those that creators of original works would in general develop or license others to develop.”\textsuperscript{88} It is unlikely that authors will license political opponents, potential critics, or would-be parodists to make any substantial use of their works.\textsuperscript{89} It is also unlikely that a market for a particular use will emerge if the costs of seeking permission outweigh the potential benefits of that use.\textsuperscript{90} In either case, because no market is likely to emerge, there is no cognizable market effect. As has already been


\textsuperscript{85} Sag, supra note 76.

\textsuperscript{86} A&M Records v. Napster, Inc., 239 F.3d 1004, 1017 (9th Cir. Cal. 2001) (noting the copyright owner’s right to develop alternative markets).


\textsuperscript{89} Id. at 591-592.

\textsuperscript{90} Gordon, supra note 22.
alluded to, Google would argue that there could not be a market for
copyright permissions in the accumulation of data for a comprehensive
book search engine because the costs of proactively clearing rights are
simply too high.

The second limiting principle relates to the idea-expression distinction. As
the reverse engineering cases aptly demonstrate, a copyright owner has no
protectable interest in preventing the copying of unprotectable expression.\(^\text{91}\)
Google would argue that although it is copying expressive works to build its
library database, its use of those works is entirely non-expressive in the
capture and processing phases—the works are not copied for their
expressive content, but rather to create a vast catalogue of metadata which
enables the search engine to operate. As I have argued elsewhere, this kind
of “non-expressive use” may well constitute its own distinct category of fair
use.\(^\text{92}\)

The relevant third limiting principal is that transformative uses do not
usually have an effect on either the market for, or the value of, the
copyrighted work.\(^\text{93}\) Google would argue that its use is highly
transformative because it provides an entirely new use for the original
works and because it is transforming concrete text into abstract metadata.
As the Second Circuit noted in the \textit{Bill Graham} case, copyright owners may
not preempt exploitation of transformative markets by simply charging
licenses for what would otherwise be fair use.\(^\text{94}\)

Even assuming that the court concluded that the scanning for inclusion in a
book search database was fairly within the copyright owners’ notional
market, the court would still need to consider the implications of the opt-out
mechanism Google has promised to comply with. If authors do have the
right to determine whether their works become part of the book search
engine, then the value of that right for each author equates to the
compensation the author could demand in return for not withholding his or
her works. For an individual book, the value of the right is determined by
the contribution that book makes to the search engine as a whole. With this

\(^{\text{91}}\) Sag, \textit{supra} note 76.
\(^{\text{92}}\) Sag, \textit{supra} note 62.
\(^{\text{93}}\) \textit{Castle Rock Entertainment v. Carol Publ'g Group}, 150 F.3d 132, 145 (2d Cir. N.Y. 1998)
(citing \textit{Campbell}, 510 U.S. at 591) “the more transformative the secondary use, the less
likelihood that the secondary work substitutes for the original.” See also, \textit{Bill Graham
Archives v. Dorling Kindersley Ltd.}, 448 F.3d 605, 614-15 (2d Cir. 2006).
\(^{\text{94}}\) \textit{Bill Graham Archives v. Dorling Kindersley Ltd.}, 448 F.3d 605, 615 (2d Cir. 2006).
Note that one can accept the reasoning of the \textit{Bill Graham} case without necessarily
endorsing its application to the facts.
conception of the market in mind, Google could quite plausibly maintain that a low-cost opt-out measure effectively negates any potential adverse market effect. After all, whether the system is opt-in or opt-out, the author should be able to realize the value of an individual work by offering to opt-in, or offering not to opt-out.\footnote{The assumption that the opt-out is truly low-cost in every salient dimension is essential to this conclusion.}

If a low-cost opt-out makes no difference, why would any copyright owner object to the opt-out? There are two principle reasons. First, opting out may not be a low-cost proposition after all. The process might be difficult and time-consuming; it might also lead to uncertainty as to what other systems an author or copyright owner should be opting out of. In other words, even with zero monetary costs, the opt-out might impose substantial implementation and information costs. In the context of the Google Book Settlement, many publishers have apparently complained that it is prohibitively expensive for them to determine if they actually own the rights in a book in order to be able to opt in or out.\footnote{Thanks to Peter B. Hirtle, Senior Policy Advisor, Scholarly Resources and Special Collections, Cornell University Library for this observation.}

Second, as the class-action lawsuit illustrates, an opt-in regime gives active copyright owners the power to object on behalf of all copyright owners, a power which can be used to extract rents far exceeding the value of those copyright owners’ individual contributions. The Authors Guild represents around 8000 U.S. authors and yet its objects to Google’s digitization project on behalf of millions of authors worldwide. An opt-in regime empowers copyright owners to act strategically to extract some of the surplus value in a joint enterprise contributed by authors who would most likely consent to the use of their works if asked or the independent investment of the defendant. While this might be a legitimate commercial strategy, it seems doubtful that a copyright owner’s inability to extract rents beyond the value of her own contribution is cognizable as a relevant market harm under the fourth factor. Analysis of the fourth factor is complicated, but the balance of the argument seems to way in Google’s favor.

5. The Likely Fair Use Outcome

Fair use is a flexible doctrine and the courts have shown a willingness to use it adaptively in the past to enable the emergence of new technology. Google has a number of strong fair use arguments within the traditional four factor framework surveyed above. Google’s position becomes stronger on the assumption that any fair use ruling would be conditioned upon the
effective provision of a relatively low-cost opt-out mechanism for copyright owners who do not want their works to be part of the Google Book database. It seems likely that if the court were to rule in Google’s favor, that ruling be subject to this condition. Of course, the court could simply rule that digitizing library books in order to extract metadata constitutes fair use with or without the opt-out—but a narrower ruling confined to the specific facts of the case is usually more likely than a broad one.

No matter how strong Google’s fair use argument may have been, the truism remains that the only certainty in litigation is that it is expensive. Google was risking a substantial possibility of statutory damages in addition to its sunk costs in relation to the GBS project. In addition, an adverse precedent in relation to book search had the potential to undermine other key legal precedents in Google’s favor with respect to its core search engine business. In short, regardless of the strength of Google’s fair use argument there was a sound business case for settling with the plaintiff authors and publishers.

PART II AN OVERVIEW OF THE GOOGLE BOOK SETTLEMENT AGREEMENT

The proposed Settlement is a complicated document consisting of 141 pages, 160 definitions, 17 separate articles and 116 separate clauses, not including the substantial provisions contained within the 15 attachments. Any attempt to summarize such an agreement in short compass is necessarily incomplete, but the following sections attempt to at least sketch out the most significant aspects of the agreement.

Google makes several significant financial concessions under the Settlement, including one-time payments of over $100 million dollars and a revenue sharing agreement. In return for these concessions Google receives the right to continue to operate its book search engine, substantially in its current form. Google will also be able to explore new revenue possibilities in cooperation with authors and publishers. The Settlement provides for

---

97 Mark Lemley commented to the Wall Street Journal: “The consequences of a loss for Google are enormous. If the publishers were to actually prevail in this lawsuit, I think it would be essentially impossible to maintain a search engine.” [citation] Mike Madison “This is not only bet-the-company litigation, it’s bet-the-Internet litigation.” Google Print II, Posted by Mike Madison - October 20th, 2005 at http://madisonian.net/2005/10/20/google-print-ii/.

98 This preliminary summary of the operative terms of the Settlement is not exhaustive and is no substitute for individually tailored legal advice. This Article does not constitute legal advice, nor does it establish any to any attorney client relationship.
extensive book previews to be made available through Google Book, it also envisages new products and services such as consumer electronic book purchases, and institutional subscriptions and various other “New Revenue Opportunities” that remain subject to negotiation between Google and the rightsholders. 99 These new features take the database well beyond the realm mere indexing and identification which enhances the value of paper collections. Some of these revenue models will directly substitute for paper-based mediums.

On of the key features of the agreement is the creation of new institutional arrangements governing the relationship between Google and the class-members. These new institutional arrangements are divided between the mechanics of the agreement itself, the newly created “Book Rights Registry” and the “Author-Publisher Procedures”. The main agreement contains a number of provisions establishing procedures and default rules governing Google’s use of copyrighted works within the Google Book universe of products and services. The creation of the Registry is envisaged as a free-standing entity which will allow participating rightsholders to collectively exploit their works and receive payment for various uses made by Google. As well as facilitating the actions of active rightsholders, the Registry will stand in the shoes of inactive rightsholders and make certain decisions on their behalf. Completing this new institutional triangle, the Author-Publisher Procedures will determine in many cases who controls the exploitation of a work within the Google Book universe and who benefits from that exploitation.100

The treatment of orphan works pervades all aspects of the Settlement. The Settlement increase public access to orphan works by presumptively including almost all works in most commercially significant uses. Orphan works will be digitized, indexed, made available for partial-previews, sold as consumer purchases and incorporated into institutional subscriptions. As well as benefiting Google, revenues attributable to these uses will flow in part to the Registry, and to registered authors and publishers.

A. The Scope of the Settlement

The effect of the Settlement is sweeping. The plaintiffs in the underlying lawsuits represent only a small section of the broader author and publisher communities, yet the Settlement embraces the vast majority of books published in the U.S. between 1923 and 2009. More precisely, the

99 See infra, Part II.
100 See infra, Part II-F.
The settlement class consists of rightsholders who have a copyright interest in any book registered with the U.S. Copyright Office and published before January 5, 2009, which is the date of the first notice of the class action Settlement. The class also includes foreign rightsholders with a copyright interest in unregistered works published in the relevant period. The definition of books covered by the Settlement expressly excludes works such as periodicals, personal papers, sheet music and songbooks, and government works (which are not protected by copyright law in any event). The Settlement also covers independently copyrighted works within books which are defined as “inserts”. Inserts include forewords, afterwords, prologues, epilogues, essays, poems, quotations, letters, song lyrics, or excerpts from other books, periodicals or other works. The term insert is specifically defined in the agreement to include children’s book illustrations, musical notation (i.e., notes on a staff or tablature) and tables, charts and graphs but to exclude, illustrations other than children’s book illustrations and pictorial works generally, such as photographs, maps or paintings. These inserts are covered by the Settlement agreement only to the extent they are independently protected by the Copyright Act. A copyrighted insert in an otherwise public domain book is covered under the agreement.

Public domain works are not affected by the agreement. In spite of this apparent breadth, two significant limitations are worth noting. First, the Settlement only covers the past, Google will be required to negotiate with authors and publishers for all future works. Second, the Settlement does not release Google from any liability with respect to foreign copyrights and Google may still be liable in other jurisdictions in which the Google book search engine is accessible. It is also important to note that many foreign authors may not realize that their U.S. copyrights are affected by the Settlement.

The Settlement gives Google the right to continue to operate its book search program along its current lines, but with some important modifications.

---

101 § 1.16(a)-(c).
102 Id.
103 § 1.16(i)-(v).
104 § 1.72.
105 § 1.72.
106 § 1.72.
107 § 1.72(b).
108 § 1.16, 2.2(i). (The Settlement neither authorizes nor prohibits, nor releases any claims with respect to the use of any work or material that is in the public domain under the Copyright Act.)
Prior to the Settlement Google allowed minimal access to books except where the book was in the public domain or was subject to a specific agreement with one of Google’s “publisher partners”.\textsuperscript{109} This resulted in three basic tiers of access: a “snippet” view by default; various levels of extended preview determined by agreement with the relevant copyright owner; and full access for public domain works.\textsuperscript{110} The same basic tiers of access remain under the Settlement with the addition of a “no-display” option.\textsuperscript{111} In addition, the Settlement envisages new services including a “Consumer Purchase” model, “Institutional Subscriptions” and various other “New Revenue Opportunities” that remain subject to negotiation between Google and the rightsholders. The Settlement gives rightsholders substantial individual flexibility to determine their own preferred level of access to their works as well as the ability to effectively opt-out of the GBS altogether. On the other hand, the Settlement sets many of the default positions very favorably for Google, and thus effectively gives it a broad license from the otherwise silent owners of orphan works.

\textbf{B. The Registry}

One of the most important features of the Settlement is the creation of the “Registry”.\textsuperscript{112} The Registry acts as both an information intermediary and a collecting society. In a letter to members, Author’s Guild President Roy Blount Jr., referred to the Registry as “the writer’s equivalent of ASCAP.”\textsuperscript{113} However, the powers of the Registry will extend far beyond those of a typical copyright collecting society.

The Registry envisaged under the agreement will own and maintain a rights information database linking books to their appropriate authors and publishers.\textsuperscript{114} The Registry receives considerable assistance from Google in this regard. Under the agreement, Google must supply to the Registry, and update regularly, a list of all books that it has digitized, the name of any library to which it has given a digitized copy of a book, a list of books sold, usage data, claims process data, library scans, and any additional

\begin{flushleft}
\textsuperscript{110} Id.
\textsuperscript{111} § 4.3(a)-(b).
\textsuperscript{112} See, Article VI — Establishment and Charter of Registry.
\textsuperscript{114} § 6.1(b).
\end{flushleft}
information necessary for the Registry to perform its obligations. In return, the Registry must provide to Google a list of all books claimed by a rightsholder, and a list of all registered rightsholders. The Registry and Google will provide each other with information bearing on the copyright status of works.

The Registry is charged with the mission of locating rightsholders, receiving payments on their behalf and distributing those payments in accordance with the detailed procedures set up under the Settlement Agreement. Consequently, as the Registry locates and identifies rightsholders it will reduce the number of works which are effectively orphaned. Some might question whether the Registry has the correct incentives to do this, however it is clearly part of the Registry’s mission. Further class-action lawsuits on behalf of the owners of orphan works may result if the registry drags its heels in this regard. The Registry must promptly distribute funds to known rightsholders as it receives them. In addition, the Registry will hold unclaimed funds for a period of 5 years, after which these funds can be used to defray operational expenses of the Registry, and any remaining unclaimed funds after that will be dispersed on a proportional basis to rightsholders. In addition, the Registry will also play a key role in negotiating prices for the “New Revenue Opportunities”. The Registry will also have an important role to play in resolving disputes between rightsholders.

The Registry plays an important role in the process of determining which books should be classified as public domain, commercially available, or in-print/out-of-print. These classifications control how Google can use the material and how conflicts between authors and publishers should be resolved. The Registry is responsible for communicating expressed rightsholder preferences to Google and, in certain instances, for standing in

---

115 § 6.6(a). NB: Google will also pay $34.5 million to the Registry, in order to fund its start up and provide for its operational costs. §5.2, and § 6.4(a). In addition, Google will provide technical assistance with respect to the design, development, and maintenance of the Registry. § 6.4(b).
116 § 6.6(c)(i), (ii).
117 § 6.6(d).
118 § 6.1(c), (d).
119 § 6.1(c).
120 § 6.1(d).
121 § 6.3(a)(i).
122 § 4.7.
123 § 6.1(e).
124 § 3.2.
the shoes of rightsholders who express no preferences.\textsuperscript{125} For example, under Section 3.2 of the Agreement, the Registry has the power to change the classification of a book or group of books from “No Display” to Display” as long as the relevant rightsholders have not indicated a preference to the contrary.\textsuperscript{126}

The powers of this new Registry are quite broad. The Registry must be authorized to act on behalf of rightsholders as set forth in the Settlement Agreement,\textsuperscript{127} but the Registry will also have other responsibilities articulated in the Settlement Agreement and its (as yet unwritten) Charter.\textsuperscript{128} The Settlement also states that the Registry is to have “such other responsibilities … as the Board of Directors of the Registry may determine are consistent with the Registry’s functions and are not inconsistent with this Settlement Agreement.”\textsuperscript{129}

The Registry’s powers are subject to some important limitations designed to protect the interests of minority class members and Google. The author subclass and publisher subclasses will be represented equally on the Registry’s board of directors to ensure a balance between individual authors and publishing companies.\textsuperscript{130} It is unlikely however that anyone on the board will be directly representative of the owners of orphan works. Furthermore, each act of the Board requires a majority of the Directors, including at least one director who is a representative of the Author subclass, and one director who is a representative of the publisher subclass.\textsuperscript{131} The Registry is allowed to represent the interests of either all authors or all publishers as a sub-group, however, the Registry is prohibited from representing any further particularized subgroups regarding any matter under the Settlement Agreement. The Registry is expressly banned from working with any sub-group of rightsholders to exclude their books from particular uses or advocating that any sub-group of rightsholders decrease its participation in the Settlement in any manner.\textsuperscript{132}

The Registry plays a significant role in relation to orphan works. Some commentators have observed that whereas a fair use ruling would have

\textsuperscript{125} § 3.2(d), 3.2(e).

\textsuperscript{126} §3.2(e). (But note that a direction to change the classification from “Display” to “No Display” can only be made by the rightsholder.)

\textsuperscript{127} § 6.1(a).

\textsuperscript{128} § 6.1(f).

\textsuperscript{129} § 6.1(f).

\textsuperscript{130} § 6.2(b).

\textsuperscript{131} § 6.2(b).

\textsuperscript{132} § 6.2(b).
substantially mitigated the orphan works problem for all potential entrants into the field of book search, the Settlement agreement only solves that problem for Google, and thus effectively creates an orphan works monopoly. For example, Pamela Samuelson argues that “the settlement would, in effect, give Google the exclusive right to commercially exploit millions of orphan books.” The Internet Archive also protests that the Settlement “effectively limits the liability for the identified uses of orphan works of one party alone, Google Inc.” James Grimmelmann concludes that “Google’s extraordinary market power under the settlement will come from its unique lock on orphan works.”

The agreement contemplates that the Registry may at some stage have the power to negotiate licenses of orphan works to third parties; however the agreement does not contain an express grant of authority from the class members to Registry to that effect. Section 6.2(b), which addresses the organizational structure of the Registry, provides that “[t]he Registry will be organized on a basis that allows the Registry, … to the extent permitted by law, license Rightsholders’ U.S. copyrights to third parties.” Section 3.8 which deals with the effect of other agreements and changes in law, contains a most favored nation clause that specifically provides that the Registry must grant Google terms that are equally favorable to any authorizations to third parties with respect to the rights of a significant portion unregistered rightsholders. In the absence of an express grant of authority from the class members to the Registry, these provisions do not appear to give the Registry the necessary authority to license orphan works to third parties. The Settlement does however leave the door open for the Registry to exercise such powers if they are conferred through some other source, such as future orphan-works legislation.

134 Samuelson, supra note 3.
136 Grimmelmann, supra note 3.
137 § 6.2(b).
138 This MFN provision only applies to authorizations made within 10 years of the Effective Date of the Settlement Agreement.
139 Indeed, it is possible that such an authorization would be beyond the legitimate scope of a class-action settlement.
Based on this interpretation, and in the absence of supervening orphan works legislation, Google will have the unique ability to commercialize orphan works, i.e. works which remain unregistered with the Registry. What remains to be seen is whether Google’s unique access to these orphan works rises to the level of a barrier to entry in any relevant market. To begin with, the economic significance of the alleged orphan works monopoly is quite speculative. Furthermore, whatever that significance is now, it should be declining. The Settlement and the establishment of the Registry provide an incentive and a means for copyright owners to reclaim active ownership of their works—thus effectively de-orphaning them. Finally, although the Registry will probably not have the right to license orphan works to third parties, it will still be in a position to manage the rights of a large number of active authors and copyright owners.

C. Financial Terms of the Settlement

The immediate cost of the Settlement for Google will exceed $100 million. Google will transfer a minimum of settlement fund of $45 million to the Registry to be distributed among the class members whose books have been digitized prior to the Opt-Out Deadline.\(^\text{140}\) Class members will receive at least $60 per book and $15 per Insert.\(^\text{141}\) Google will also pay $34.5 million to fund the launch and the initial operations of the Registry and $30 million to pay the plaintiffs’ attorneys fees and costs.\(^\text{142}\) Thus, Google will pay at least $109.5 million under the Settlement agreement regardless of the revenue it generates from its book related products and services authorized under the agreement.

On first impressions, Google has agreed to a 30/70 revenue split with rightsholders for advertising and purchase revenues attributable to works covered by the agreement.\(^\text{143}\) The agreement entitles rightsholders to 70% of “Net Advertising Revenues” and “Net Purchasing Revenues”, however, these terms are defined with a built-in 10% deduction for Google’s operating costs.\(^\text{144}\) This 10% deduction can itself be offset pursuant to the unclaimed funds provision.\(^\text{145}\) Which means that, in theory at least, active authors can profit from the exploitation of orphan works and misclassified public domain works. It is important to note that the revenue split does not

\(^{140}\) § 2.1(b). ("Opt-Out Deadline" means the deadline fixed by the Court to opt-out of the Settlement pursuant to Rule 23 of the Federal Rules of Civil Procedure.)

\(^{141}\) § 5.1(a).

\(^{142}\) § 5.1(c), 5.5.

\(^{143}\) § 4.5(a)(i)-(ii).

\(^{144}\) § 1.86, 1.87.

\(^{145}\) § 6.3(a).
apply to works outside the scope of the Settlement, such as public domain works. Nor does it apply to initial search result pages which feature more than one work.\textsuperscript{146}

\textit{D. Uses and Permissions Under the Settlement}

1. Book Search: Digitization & Display

The Settlement avoids resolving whether the digitization and non-expressive use of copyrighted books is permissible under copyright law. Instead, the Settlement grants Google a very broad license in relation to its book search program. In several key aspects, the uses Google is permitted to make of these works goes well beyond the best plausible fair use ruling Google could have hoped for. In other many respects, however, the terms of the Settlement agreement closely mirror the likely outcome of law suits which the Settlement averted.

The most important way in which the Settlement mirrors the likely fair use ruling is in the centrality of the copyright owners’ right to opt-out. As discussed previously, the most likely result of the copyright litigation avoided by the Settlement would be a finding in favor of fair use, conditional on the effective provision of a relatively low-cost opt-out mechanism. The ability of rightsholders to opt-out is also a key feature of the Settlement. Not only can copyright owners opt-out of the class action Settlement altogether,\textsuperscript{147} under the Settlement itself they can opt-out of inclusion in the book search database or particular elements thereof.\textsuperscript{148} As detailed below, the Settlement gives individual rightsholders substantial flexibility to control the use of their works by Google, to determine their preferred level of public access, and in some cases, to control pricing. The division of benefits and control between authors and publishers is governed by the Author-Publisher Procedures.\textsuperscript{149}

The agreement authorizes Google to digitize books and inserts obtained by Google from any source and to use that digitized information in accordance with the agreement. This license applies to all books and inserts covered by

\textsuperscript{146} Google does not currently advertise on these menu pages.
\textsuperscript{147} § 17.33.
\textsuperscript{148} A rightsholder may choose to opt-out of the Settlement prior to the “Opt-Out Deadline”, the deadline fixed by the Court to opt-out of the Settlement pursuant to Rule 23 of the Federal Rules of Civil Procedure. § 17.33. The Settlement does not authorize or permit the use of books for which all rightsholders opt-out. \textit{Id.} The Settlement also does not release any claims for such books. \textit{Id.}
\textsuperscript{149} See infra, Part II-F.
the Settlement agreement, however, Section 3.5(a)(i) provides that for at least 27 months after the Settlement is finalized, any rightsholder may direct Google to remove his or her book from the GBS database.\textsuperscript{150}

With the exception of works that have been removed pursuant to Section 3.5(a)(i), all covered works are subject to digitization and may be listed in an initial search results page. The extent of the information accompanying that listing depends on a host of factors beginning with the Display/No-Display classification.

Much of the agreement revolves around the distinction between “Display” and “Non-Display” uses and the accompanying categorization of books as either “Display” or “No Display”. Non-Display Uses are uses that do not display expression from digital copies of books or inserts to the public.\textsuperscript{151} The display of bibliographic information and important data about a book, such as which pages match a particular search query, the book’s key terms and chapter headings, library collection information, are all considered to be “Non-Display” uses.\textsuperscript{152} In addition, the use of digital copies of books for Google’s internal research and development purposes is also considered to be a non-display use.\textsuperscript{153} Thus, whether a book title appears as part of the initial menu of search results responding to a user query does not depend on its classification as either “Display” or “No Display”, but, critically, the extent of information accompanying the book’s basic bibliographic

\textsuperscript{150} Section 3.5(a)(i) provides that, for 27 months after the Notice Commencement Date (which will most likely be the date the settlement is approved), Google will respond to requests for removal by making “reasonable efforts” not to digitize such books and by removing which them in the event that they are digitized. After the initial 27 month period, Google will continue to making “reasonable efforts” not to digitize books subject to a removal request, but will not remove those which have already been digitized. Nonetheless, the rightsholder retains the right to request exclusion from particular Display Uses. § 3.5(a)(iii).

\textsuperscript{151} § 1.91. It is worth noting here that “Expression” is defined in the agreement as anywhere between 3 and 10 contiguous words in which a person has a copyright interest under Section 106 of the Copyright Act, without regard to the fair use doctrine and other statutory defenses embodied in Sections 107 through 122 of the Copyright Act. “Expression” remains subject to the limitation of substantial similarity (an implicit limit on Section 106) and the idea expression distinction as embodied in Section 102(b) of the Copyright Act. Technically, it is open for Google to argue that displaying a snippet that is non-infringing because it is not a substantial copy or because of the idea expression distinction falls within the scope of “non-display uses” under the agreement. However, the structure of the agreement indicates that the parties intended for snippets to be covered as a “display use”.

\textsuperscript{152} § 1.91.

\textsuperscript{153} § 1.91.
information does. If a book has been classified as a “Display” book, its listing will also be accompanied by a brief “snippet” of about three to four lines of text from the book.

The distinction between “Display” and “No Display” works becomes more apparent on the individual “Online Book Pages”. Once a user selects a particular book as the object of their investigation, they will be directed to an Online Book Page dedicated to that book. In the case of No Display books, the Online Book Page will simply contain bibliographic information and data about a book consistent with the definition of Non-Display Uses. The Online Book Page for a Display Book will feature up to three “snippets” of three or four lines of text, per search term, per user. In general, the Display Books will also permit an extended preview of the work.

The Settlement sets out three different types of “Preview” modes applicable to books. If a work is subject to the “Standard Preview”, the pages selected for display will depend on the relevant user query. The default standard preview is limited to no more than 20% of the books pages and no more than five adjacent pages for non-fiction works, not including the book’s table of contents, title page, copyright page, or index. The default page limits for works of fiction is set at 5% of the work and no more than 15 adjacent pages with the additional stipulation that the final 5% and the final 15 pages of the work must be blocked.

These are merely the default settings. Under Section 3.5(b)(i), a rightsholder may exclude a book or group of books from one or more of the display uses. In addition, Section 3.2(e)(i) allows the rightsholder to change the classification of a book or group of books from Display to No Display at any time, which amounts to a complete opt-out from the all Display Uses. Rightsholders also have the option of selecting a Fixed Preview or a

---

154 § 1.91, 3.4(a).
155 An Online Book Page is any page dedicated to a single book and advertising on such pages is subject to the revenue split.
156 § 1.97 (defining Online Book Page); § 1.147 (defining Snippet Display).
157 § 4.3(b)(i)(1). (The preview must omit at least two pages before and after the relevant set of adjacent pages).
158 Also note that Unless approved by the Registry or the Rightsholder, Google will not offer to users copy/paste, print or Book Annotation functionalities as part of Preview Uses. (4.3(b)(3)).
159 In case there was any doubt, section 4.3(a) also provides that a rightsholder may directs Google or the Registry not to allow previews.
Continuous Preview as an alternative. Under a “Continuous Preview” mode, Google is authorized to display up to 10% of pages without the adjacent page limitations of Standard Preview. Under the “Fixed Preview” mode, an initial selection of no more than 10% of the work is made available for preview, and this selection will not vary from user to user depending on the search being performed. In all three modes, Standard, Continuous and Fixed, a rightsholder may choose to increase the percentage of the work made available up to 100% of the work.

2. Beyond Book Search

The Settlement also envisages a number of extensions to the product and service offerings of the GBS including consumer purchases of electronic books, institutional subscriptions for corporations and universities and a number of other “New Revenue Opportunities”.

The agreement allows Google to offer access to Display books to the general public as “Consumer Purchases.” This system would allow users to view, copy & paste, and print books within certain limits. The ability to copy/paste is limited to no more than four pages with a single copy/paste command. Printing would be limited no more than 20 pages per print command. Although it is notable that the agreement does not appear to place any obligation on Google to prevent an entire book being printed in 20-page increments. Under the Consumer Purchase model, users will be able to make book annotations for their personal use and to share those annotations with a limited number of persons, subject to limitations. As a security measure, printed pages will include a copyright watermark.

160 § 4.3(c).
161 § 4.3(c)(iii).
162 § 4.3(c)(iii).
163 § 4.3(b)(ii). The agreement simply states that “Rightsholders may increase the percentages in Fixed Preview, Standard Preview and Continuous Preview”. Presumably an increase to 100% negates the adjacent page restrictions. Note that unless the rightsholder has authorized display of one hundred percent (100%) of the pages of a Book in Preview Use, Google is obliged to block access to at least five percent (5%) of the pages of a Book. § 4.3(d).
164 § 4.2.
165 The Security Standard does require Google to monitor for and take steps to prevent simultaneous logins from different locations using the same account. Attachment D: Security Standard, § 3.9.2.
166 § 3.10(c)(ii)(5). Annotations are not be shared with the general public. § 3.10(c)(ii)(5)(a). Annotations will only be accessible to persons authorized to view the book, and will be limited to 25 persons, all of whom must be identified individually by the user. § 3.10(c)(ii)(5)(a)-(d).
The Consumer Purchase model envisages two pricing options: either a specified price determined by the rightsholder or a “Settlement Controlled Price” determined by a pricing algorithm developed by Google. The pricing algorithm will be designed to maximize the revenue accrued to each individual rightsholder, not the total Google’s revenue. Rightsholders may also change their choice of pricing option at any time. The agreement defines a set of initial pricing bins: $1.99, $2.99, $3.99, $4.99, $5.99, $6.99, $7.99, $8.99, $9.99, $14.99, $19.99 and $29.99—but these are subject to renegotiation by Google and the Registry at any time. The allocation of works for Consumer Purchase to these initial pricing bins is illustrates in Figure 1 below.

*Figure 1. Consumer Purchase Pricing Bin Allocation*

---

167 § 4.2(b)(i)(1)-(2).
168 § 4.2(b)(i)(2).
169 § 4.2(b)(ii).
170 § 4.2(c)(i). The agreement also sets out procedures for renegotiating the number and content of the pricing bins after three years and then periodically every four years.
The agreement contemplates offering schools, universities and corporations the option of an Institutional Subscription to Google’s catalog of digital books. These Institutional Subscriptions will allow online access to and viewing of the full contents of all Display Books in the Google book database, unless of course the rightsholder opts out of that particular use. Institutional Subscriptions will enable users to view, copy/paste, and print pages of a Book, and may enable Book Annotations. Like the Consumer Purchase model, the copy & paste functions are to be limited to up to four pages with a single copy & paste command. Printing is to be limited to no more than 20 pages with one print command. The shared annotation feature will operate differently for Institutional Subscriptions than for Consumer Purchases. Here, users will be limited to sharing with instructors and students in a single academic course, in connection with such course during an academic year, or with students of the same course in subsequent academic years. Employees of the subscriber will also be allowed to use the annotations in connection with a discrete work project during the course of that project.

The financial terms of the proposed Institutional Subscriptions have already proven controversial. Robert Darnton, the Director of the Harvard University Library, argues that without competition, Google’s Institutional Subscriptions will follow the path of academic journals:

Once a university library subscribed, the students and professors came to expect an uninterrupted flow of issues. The price could be ratcheted up without causing cancellations, because the libraries paid for the subscriptions and the professors did not. … Owing to the skyrocketing cost of serials, libraries that used to spend 50 percent of their acquisitions budget on monographs now

---

171 § 4.1(d).
172 § 4.1(d).
173 § 4.1(d).
174 § 4.1(d). Like the Consumer Purchase model, Google will include a visible watermark on pages printed from the Institutional Subscription Database. The watermark will identify the material as copyrighted and display encrypted session identifying information which can be used to identify the authorized user that printed the material or the access point from which the material was printed. These watermarks will not obscure the content of the printed pages. Again, there does not appear to be any obligation on Google’s part, to prevent users from printing entire books in 20-page increments. The Security Standard does require Google to take commercially reasonable steps to verify that users are affiliated with an institution with an active subscription. Attachment D: Security Standard, § 3.9.3.
175 § 3.10(c)(ii)(5)(e).
176 § 3.10(c)(ii)(5)(e).
spend 25 percent or less. University presses, which depend on sales to libraries, cannot cover their costs by publishing monographs. And young scholars who depend on publishing to advance their careers are now in danger of perishing.\footnote{Robert Darnton, \textit{Google & the Future of Books}, February 12, 2009, http://www.nybooks.com/articles/22281.}

A chilling prospect indeed.

Professor Darnton will not find much comfort in the terms of the Settlement. The agreement provides that the pricing strategy for Institutional Subscriptions “will be governed by two objectives: (1) the realization of revenue at market rates for each Book and license on behalf of Rightsholders and (2) the realization of broad access to the Books by the public, including institutions of higher education.”\footnote{§ 4.1(a)(i).} The agreement further explicates: “Plaintiffs and Google view these two objectives as compatible, and agree that these objectives will help assure both long-term revenue to the Rightsholders and accessibility of the Books to the public.”\footnote{Id.} This, of course, makes no sense. These twin objectives—maximizing revenue and maximizing public access—can only be consistent up to a point. At the end of the day, Google and the Registry will have to choose one or the other as the primary objective.

It seems clear from the agreement that Google plans to divide up the market into segments to facilitate price discrimination. Pricing for Institutional Subscriptions will be determined in reference to: comparable third party products and services, the scope of books available, the quality of the scan and the features offered as part of the Institutional Subscription. Subscription prices will also vary depending on the size and nature of the institution (calculated on a full-time equivalent basis) and depending on the nature of the institution.\footnote{§ 4.1(a)(iii-iv).} Pricing is likely to vary across broad categories such as Corporate, Higher Education (with multiple sub-divisions based on the Carnegie Classifications for Institutions of Higher Education within the United States); School (K-12); Government; and Public.\footnote{§ 4.1(a)(iv).} The agreement prohibits remote access without Registry approval for these last three categories. As the Settlement itself notes, segmenting the market in this way will make it easier to reconcile the conflicting objectives of maximizing...
revenue and maximizing public access. The agreement also contemplates a different form of price discrimination based on offering specialized versions of the institutional subscription database. Google is required to ensure that the price for access to the entire Institutional Subscription Database will be less than the sum of the prices for access to any discipline-based specialized collections.

The Settlement allows Google to provide not-for-profit Higher Education Institutions and Public Libraries with limited free access under the “Public Access Service.” This free public access is limited to one terminal per library building for libraries, one terminal per 4000 students for schools qualifying as Associate’s Colleges, and one terminal per 10,000 for all other schools. The Registry and Google may also agree at some future date to make additional terminals available, either for free or for an annual fee.

In addition to viewing these works electronically, the agreement also provides that page by page printing may be allowed from these terminals at a per-page fee to be determined by the Registry. A feature of the agreement that will no doubt be of interest to universities, copy shops and copyright collecting societies is the potential to directly link the Google book database to copy shops. The Settlement provides that the Registry and Google may agree to make a commercial public access service available to copy shops and other entities for an annual fee per concurrent user and a fee per printed page. It is too early to tell whether this might lead to a reduction in the cost of course packets.

§ 4.1(a)(v).

§ 4.8(a)(i).

§ 4.8(a)(i)(2).

§ 4.8(a)(ii).

§ 4.8(a)(iii).

§ 4.8(b).

§ 4.8(b).

§ 4.8(b).

§ 4.8(b).

§ 4.8(b).

§ 4.8(b).
3. The Research Corpus

The Settlement provides for the creation of a Research Corpus for non-consumptive and non-commercial research by certain qualified users. The Research Corpus consists of all Digital Copies of Books made in connection with the Google Library Project, other than Digital Copies of Books that have been removed pursuant to Section 3.5 or withdrawn pursuant to Section 7.2(d)(iv). A Rightsholder may withdraw its book from the research corpus if the book is commercially available as of the Notice Commencement Date, or becomes commercially available within two years of the Notice Commencement Date. However, the book will be returned to the research corpus if it ceases to be commercially available after that two year period. Books removed pursuant to Section 3.5 are also withdrawn from the Research Corpus.

Use of the Research Corpus is limited to “Non-Consumptive Research”: i.e. research involving computational analysis of a book or a set of books but which does not involve the use or display of substantial portions of a book to understand the intellectual content presented therein. The term “non-expressive use” has been used to denote the same concept.

Commercial exploitation of algorithms developed while performing non-consumptive research is permitted under the agreement. However, the use of data extracted from specific Books within the Research Corpus to provide services that compete with services offered by the rightsholder of those Books or by Google is prohibited. Direct, for profit, commercial use of information extracted from Books in the Research Corpus is prohibited without the express permission of the Registry and Google. However, researchers may report the results of their Non-Consumptive Research in scholarly publications whether these are commercial or not. The agreement contains a detailed set of provisions for evaluating research proposals and monitoring the activities of qualified researchers.

---

189 § 7.2(d)(i).
190 § 1.130.
191 § 7.2(d)(iv).
192 § 7.2(d)(iv).
193 § 7.2(d)(vi)(1).
194 Sag, supra note 62.
195 § 7.2(d)(x).
196 § 7.2(d)(ix).
197 § 7.2(d)(viii).
198 § 7.2(d)(vii).
199 § 7.2(d)(xi).
Generally, qualified researchers will be academics affiliated with libraries and universities. In addition to researchers from affiliated with libraries involved in the GBS, suitably accredited researchers affiliated with universities, United States not-for-profit research organizations (such as a museums, observatories or research laboratories), or United States governmental agencies, provided that have their research agendas approved in advance. An individual who does not meet these criteria may become a “Qualified User” by demonstrating that she has the necessary capability and resources to conduct non-consumptive research. For-profit entities may only become qualified researchers if both the Registry and Google give their prior written consent.

The non-consumptive research contemplated in the Settlement may be valuable to Google’s competitors. While competitors cannot directly extract information from books in the research corpus, they should benefit from research pertaining to the research corpus generally. While Google book search will clearly have spillover benefits for Google, it also seems likely that the research corpus has the potential to benefit both Google and its competitors. The more open access to the research corpus it is, the less weight should be given to the claimed that Google’s control of the book search database gives it an advantage over competitors in adjacent fields such as Internet search more generally.

4. New Revenue Models

The Settlement contains what amounts to an agreement to agree with respect to new revenue models, including: print on demand; custom pay per-page publishing for course materials in the educational and professional markets; Downloadable PDF versions of Books; consumer subscription models which would be an individual version of an Institutional Subscription; and products comprised of summaries, abstracts and/or compilations of books. These New Revenue Models are likely to be applied to orphan works, subject to Registry approval. The economic terms for the New Revenue Models will not necessarily adhere to 70/30 revenue split for advertising and purchases. The economic terms may be renegotiated at the option of either Google or the Registry after three years.

---

200 § 1.121 (defining Qualified User); 7.2(d)(xi)(2) (Research Agenda).
201 Id.
202 Id.
203 § 4.7.
204 § 4.7.
and then periodically every four years, or as otherwise mutually agreed by Google and the registry.\(^\text{205}\)

### E. Book Classification

The Settlement comprises a complicated collection of default settings, modification provisions and opt-out rights. For copyright owners who are actively managing their works, the most important provisions of the agreement will be those which allow them to customize how their works are used. However, the default settings embodied in the Settlement are extremely important because they control how the vast majority of works which are not being actively managed (i.e. orphan works) will be made available through the GBS. The default settings are consequential because large numbers of authors and publishers are unlikely to take advantage of the numerous opt-in and opt-out provisions sprinkled throughout the agreement. Some of the default positions are subject to modification by the Registry, however, usually the Registry only has the option to make a work more available, not less. The Settlement defaults may also be significant in setting the expectations of active rightsholders because they define a new status quo and baseline for negotiation. A number of the significant default positions and customization opportunities have been noted in the preceding analysis already. What remains to be discussed is the all-important question of books classification. Google will initially classify all books as either commercially available or not commercially available as of the Notice Commencement Date.\(^\text{206}\) This classification in turn determines, at least initially, whether a book is Display Book or a No Display Book.\(^\text{207}\) The agreement sets out the process for determining whether a book is commercially available in significant detail.\(^\text{208}\) To be commercially available, a book must be offered for sale new through one or more customary channels of trade in United States.\(^\text{209}\)

Google is required to make this commercial availability determination based on an analysis of third-party databases and information publicly available on the Internet.\(^\text{210}\) The Settlement recognizes that these sources of information may contain errors. To the extent that the information available from multiple sources is in conflict, Google is required to make its availability determination in a manner that is reasonably designed to

\(^{205}\) § 4.9.

\(^{206}\) § 3.2(b).

\(^{207}\) Id.

\(^{208}\) § 3.2(d)(i).

\(^{209}\) Id.

\(^{210}\) Id.
minimize the overall error rate. The Settlement also allows rightsholders to provide information directly to Google that a book is commercially available. Google must act on this notification within thirty days unless it reasonably believes that the information is inaccurate.211 A rightsholder can also notify Google directly or indirectly through the Registry of any mistaken classification in either direction.212 When no information is available from these sources, the book will be classified as not commercially available.213 This same initial determination of whether or not a book is Commercially Available will be used to initially classify Books as “in-print” or “out-of-print,”—a classification which can then by altered by the terms of the Author-Publisher Procedures.214

In general, books in-print that cease to be commercially available will be reclassified as out-of-print.215 If the Registry re-classifies a book as out-of-print, it will direct Google to change the book’s classification to Display Book.216 However, if the Registry reclassifies a book as in-print, the Registry will direct Google to change the classification of the book to a No Display Book only upon the express request of the Rightsholder of the book.217

These classifications are only the starting point. Regardless of the initial classification, and individual rightsholder, or the Registry, can change a

211 Id.
212 Any disputes regarding the determination of whether a Book is Commercially Available are subject to the Dispute Resolution Procedures pursuant to Article IX of the Settlement.
213 In the event that the database contains multiple versions of the same work, all versions will be treated as commercially available so long as one of those versions is subject to copyright and commercially available. Likewise, if a public domain book is commercially available with an Insert (such as a foreword or preface) which is subject to copyright protection, then any earlier edition of such Book that contains such Insert is also deemed to be Commercially Available. Note that foreign language translations, abridged versions and versions with different forewords or additional annotations are treated as separate books under the agreement. Likewise, even though a hardcover book and a soft cover book may contain the same principle work, they are nonetheless considered different books under the settlement agreement if they have different ISBNs. Section 1.111 (definition of “Principal Work”). Note also that at any time after one year from the Final Approval Date, Google shall have the right to request that the classification of a Book be changed to a Display Book if Google believes that the Book is not Commercially Available at that time, or if Google believes that a mistake was made in initially determining the Book to be Commercially Available. § 3.2(e)(ii).
214 See infra Part II-F.
215 § 3.3(a). Either an author or a publisher may challenge the classification of a book by filing a notice with the Registry APP Section 3.3 (b).
216 Subject to any Change Requests by Rightsholders under Section 3.2(e)(i).
217 Again, subject to any Change Requests by Rightsholders under Section 3.2(e)(i).
book’s classification and therefore effect the rights and permissions associated with that book. Section 3.2(e)(i) allows the rightsholder to change the classification of a book or group of books from Display to No Display at any time, which amounts to a complete opt-out from the all Display Uses.  But note that while a direction to change the classification of a book to a No Display Book must be initiated by the Rightsholder of the book, the Registry may change the classification of a book or group of books from No Display to Display on its own initiative. In addition to changing the No Display to Display classification of a book, a rightsholder may direct Google to exclude a book, or a portion thereof, from any one or more, or all display uses, revenue models, or the book annotation sharing feature. This direction may be given at any time.

The process for identifying books in the public domain is also an important part of the Settlement. The Settlement contains a safe harbor for Google with respect to books that it mistakenly classifies as public domain. To take advantage of the safe harbor, Google must follow a specific process for determining whether book is in the public domain. The process is not mandatory, but it must be followed for the safe harbor to take effect.

Under current U.S. copyright law, a work will become part of the public domain when the term of copyright protection expires, or if the work is a US government work. In addition, in the past it was possible for a work to enter the public domain because it was published without the appropriate copyright notice or because the copyright owner failed to renew his or her registration. Although these formalities and no longer a part of the current U.S. Copyright Act, they still affect the copyright status of works published in the past. The extraordinary length of copyright protection and the extraordinary legal and factual complexity of determining whether a work is in the public domain significantly diminishes the usefulness of the public domain as a cultural resource. One of the anticipated public benefits of the Settlement is that it creates strong incentives for Google to correctly identify which works are in the public domain. Again, one could question whether the Registry will be properly motivated to identify books as belonging to the public domain and thus outside their control.

---

218 § 3.2(e)(i).
219 § 3.2(e)(i).
220 § 3.5(b)(i).
221 Id.
222 Attachment E to the Settlement Agreement (Public Domain).
F. Author-Publisher Procedures

The Author-Publisher Procedures in Attachment A to the Settlement Agreement (“APP”) play a substantial role in determining who controls the exploitation of a work within the Google Book universe, and also who benefits from that exploitation. The APP defines the central question of who is regarded as a Rightsholder under the agreement; who speaks for the work with regard to opting into or out of Display Uses; who determines the price of access to the work under the Consumer Purchase model; and how revenues from electronic distribution and advertising are allocated between authors and publishers. Before reviewing these questions in detail however, it is first necessary to navigate an obstacle course of threshold conditions and definitions to which the substantive provisions of the APP are keyed.

1. In-Print versus Out-of-Print

The first determination required to make sense of the APP is the distinction between in-print and Out-of-print. Article 3 of the APP sets out the ways in which Google and the Registry will initially classify books either in-print or Out-of-print, how they will test and change that classification, given a successful challenge by either the author or publisher (or their representatives on the Board of Directors on the Registry) under two additional tests for determining a work’s in-print/out-of-print status. If a book meets either of the two tests, it will be classified in-print, otherwise it will be classified out-of-print.

Confusingly, the first reclassification test combines two distinct contractual concepts—reversion of rights to the author and the status of the book as in-print under the relevant author-publisher contract. If the author-publisher contract does not provide for reversion to the author under any

---

223 The APP do not otherwise alter the rights and obligations of the parties. APP § 2.1. The APP are effective as of the Effective Date defined in the Settlement, but that if the Registry is operational prior to that date, an Author and Publisher may agree to invoke the APP. APP § 2.2.
224 APP § 3.1.
225 APP § 3.2(a), § 3.2(b).
226 Id.
227 APP § 3.2(a)(i)
circumstances, the book is regarded as in-print regardless of whether it is commercially available or not. If the author-publisher contract does provide for reversion, the question then becomes whether the book is in-print under that contract.

This test is confusing because it implicitly assumes that all author-publisher contracts use in-print/out-of-print as the sole trigger for reversion of rights to the author. In fact, author-publisher contracts vary enormously depending on the relative bargaining strengths of the parties, the previous experiences of both the author and the era in which the contract was put into effect. To address this potential problem, the APP provides that a book will not be considered “in-print” if the contract provides for reversion to the author, and all of the criteria for reversion have been met.

Nonetheless, the problem of determining whether the book is in-print under the terms of the author-publisher contract still remains. Contracts drafted before electronic distribution and print-on-demand became significant issues are unlikely to be easy to interpret in this respect. For example, it is unclear whether a single book published and shipped upon demand by a book store means that it “available for sale either from stock, from publisher’s warehouse, or in the field”, thus making it not out-of-print according to many publishing contracts. The APP does not answer this question, but it does provide some assistance in interpreting other terms of author-publisher contracts which may be used to trigger reversion to the author. The APP provides that if the determination of in-print status under the contract is based on revenues, and more than 50% of the revenues come from the Google Book Revenue Models, then those revenues will not be used to meet the requirements. An equivalent principle applies to determinations based on units sold. The APP also provides that mere inclusion of a book in a database or search results will not provide sufficient proof that a book is “in-print.”

The second reclassification test for challenging a book’s in-print/out-of-print status relates to future publications. Specifically, the APP provides that a book will be considered in-print if, consistent with its contractual

---

228 Id.
229 Id. Note that the Book may be “in-print” even if the contract does not use the term “in-print.”
230 APP § 3.2(a)(i)(3).
231 APP § 3.2(a)(i)(1).
232 Id.
233 Id.
rights, the publisher publicly announces to the trade that it has undertaken concrete steps to publish an existing or new edition of that book, and such edition is published within twelve months of the announcement. It is unclear whether a book would be retrospectively reclassified as out-of-print if the publisher’s announced intention to publish within twelve months did not in fact come to fruition. It is also unclear how many copies must be so published, and in what form. Presumably, the announced publication would have to meet the contractual threshold determining whether a book is in-print.

2. Author-Controlled Books

In addition to the in-print/out-of-print classification discussed above, the APP contains a separate provision determining whether book is to be classified as “Author-Controlled.” As explained in more detail below, this classification is important because it allows the author alone to direct the exploitation of a book by Google: including decisions about removal, exclusion, changes in display uses, and pricing. This classification also means that the author receives 100% of the revenue from the book.

Section 4.1 of the APP sets out the procedure to be applied in determining whether any book for which the rights have not reverted to the author should nonetheless be classified as “Author-Controlled”. The procedure does not apply to a work-for-hire or a work for which the author-publisher contract does not provide for reversion to the author under any circumstances. Although this process may be broadly consistent with various author-publisher agreements, where it is not, it effectively overwrites the contractually mandated procedures in such agreements.

To obtain the “Author-Controlled” classification with respect to a book, an author must first have sent a request for reversion to the publisher in accordance with the relevant author-publisher contract. If the publisher does not respond or did not respond within (i) 90 days or (ii) the applicable period defined by the contract, the author may send a “Change Status Notice” to the Registry, with a copy to the publisher. That notice must affirm that the author believes the book is no longer in-print, that the response period has expired, and the author must include a copy of the previously-sent request for revision and applicable author-publisher

---

234 APP § 3.2(a)(ii).
235 APP § 4.
236 APP § 4.1.
237 APP § 4.1(a).
238 APP § 4.1(b).
contract.\footnote{APP § 4.1(b)(i)-(iii).} It appears that if the relevant contract gives the publisher a “reasonable period” (or some other unspecified period) to respond, an author could rely on the specification of 90 days in the APP as defining that previously undefined period.

When the Registry receives an author’s notice, it must escrow all subsequent revenues from the book.\footnote{APP § 4.1(c).} If the publisher does not respond to the Change Status Notice within 120 days, the Registry will deem the book to be Author-Controlled.\footnote{APP § 4.1(d).} If the publisher disputes the Change Status Notice within the requisite 120 days, the parties will to the Registry for decision.\footnote{APP § 4.1(e).} When reviewing the issue, either party can ask the Registry to review the in-print/Out-of-print status of the book according to Article III of the APP.\footnote{APP § 4.1(e)(i).} The Registry will also review the evidence as to whether the rights have reverted according to the terms of the author-publisher contract, and if the conditions for reversion have been met, deem the book author controlled.\footnote{APP § 4.1(e)(ii)-(iii).} The Registry will distribute the escrowed revenues according to its determinations on the issues of Author-Control and the in-print/Out-of-print status of the book.\footnote{APP § 4.1(e)(iii)-(v).}

Article IV of the APP gives authors an alternative mechanism under which to exercise their contractual rights of reversion. Prior to the Settlement an author would have had to rely solely on his or her contractual remedies, under the APP authors now have the option of sending a petition through the Registry. If the publisher fails to respond, the author automatically prevails.

3. Who Controls the Exploitation of the Work?

Having surveyed the nuances of the in-print/ out-of-print classification and the separate, but related, classification of “Author-Controlled” works, it is now possible to address the substantive questions that arise between authors and publishers—who controls the exploitation of the work under the Settlement and who benefits from the exploitation?

\footnote{APP § 4.1(e)(ii)-(iii). If both an author and publisher claim or register a book under Section 13.1(c)(ii)(2)(B) of the Settlement, the Registry will notify both about the conflicting claims, and the protections provided under Article IV of the APP. APP § 4.2.}
The first issue is defining exactly who constitutes a Rightsholder under the agreement. This depends on whether the work is considered a work-for-hire under U.S. Copyright law, whether the book is in-print or out-of-print, whether the book is author-controlled and whether the rights have reverted to the author.

Beginning with the scenario most favorable to the author, if a work is out-of-print and the rights have reverted to the author, the author will be considered the sole rightsholder. In addition, if the work is deemed to be author-controlled (which means that it should be out-of-print), the author will be considered the sole Rightsholder. If the rights in an out-of-print book have not reverted to the author, and the work is not deemed to be Author-Controlled under Article IV, both the author and the publisher shall be considered rightsholders. For books in-print, both the publisher and the author are considered to be rightsholders under the terms of the Settlement. The only way the publisher gets to be the sole Rightsholder is if the book in question qualified as work-for-hire under U.S. copyright law at the time it was made.

The APP outlines a procedure for resolving conflicting instructions between joint-rightsholders and these procedures vary depending on whether the book is in-print or out-of-print. The upshot of these procedures is that in the event of a conflict over permissions and levels of access, the most restrictive direction will prevail. Specifically, the APP assumes that publishers will make the first move in defining the acceptable uses of books in-print. Authors are then given a 30 day window to object to some or all of the display uses requested by the publisher. If an author fails to respond within 30 days, the display uses requested by the publisher will prevail subject to Section 3.5 of the Settlement Agreement. This is an important caveat, Section 3.5 of the Settlement Agreement allows Rightsholders, at any time, to exclude a book, or part of a book, from any one or more, or all display uses, revenue models, or the book annotation sharing feature. Section 5.3 of the APP provides that any time the author and the publisher

---

246 APP § 6.1(b).
247 APP § 6.1(a).
248 APP § 6.1(c).
249 APP § 6.1.
250 APP § 5.1(a), APP § 6.1(a). If a Book is a work-for-hire, only the Publisher is considered a rightsholder for purposes of directing removal, exclusion, changes in display use, or level of access for any revenue model.
251 APP § 5.3, APP § 6.1(c).
252 APP § 5.1(a)
253 APP § 5.1., APP § 3.5(b)(i).
of a book issue conflicting directions to the Registry for an in-print book pursuant to Section 3.5, the more restrictive directions as to removal, exclusions, changes in display uses or levels of access will control. Thus even if an author misses the 30 day response window, she still retains a substantial say in the display uses and revenue models affecting her work.

The joint control provisions for out-of-print works are not an exact mirror of the rules for works in-print. To the extent the publisher retains Rightsholder status in an out-of-print book (i.e. if the rights have not reverted to the author and the work is not Author-Controlled), both the author and the publisher are can direct removal, exclusion, changes in Display Uses and/or levels of access for any Revenue Model. However, the publisher must articulate a “good reason” for its directions, such reasons include legal, technical, editorial, policy, commercial or economic considerations. But note that reducing access to one book in order to benefit another by reducing competition is not a “good reason.” Once again, if, at any time, an author and the publisher of a book issue conflicting directions to the Registry regarding the uses authorized for such out-of-print book, the more restrictive directions as to levels of access will control.

The rules applicable to pricing also depend on the in-print/out-of-print status of the work. For books in-print, so long as the publisher and author agree to authorize one or more Display Uses, then the publisher has the right to determine the pricing of the book for Consumer Purchase. Authors cannot directly control the price of an in-print book, but they can use their power to exclude a book from one or more of the display uses to get some accommodation from publishers. While this strategy may be effective for authors who simply disagree with their publisher as to the optimal price point for a work, it will not give significant leverage to authors who would like their works to be available for free or subject to a Creative Commons license. It is also worth noting that while publishing contracts rarely give authors any input into pricing decisions, to the extent that an author has such a contractual right, the Settlement appears to vitiate it for works in-print.

If both the author and the publisher are considered rightsholders with respect to an out-of-print work, either party may set the price applicable to

\[254\text{APP }\S 6.1(c).\]
\[255\text{APP }\S 6.1(c).\]
\[256\text{Id.}\]
\[257\text{Id.}\]
\[258\text{APP }\S 5.4.\]
Consumer Purchases. The Registry will act on the most recent notification received, subject to the timely objection of the other party. In the event of such an objection, the higher of the prices specified will prevail.\textsuperscript{259} Again, this highest common denominator approach to pricing for out-of-print works may contradict an expressed term of some author-publisher contracts and is likely to frustrate authors who would like to make their works more freely available through the Google Book universe.

4. Who Benefits from the Exploitation of the Work?

The APP controls the allocation of revenues generated under the Settlement’s various revenue models between authors and publishers.\textsuperscript{260} This includes the initial cash payments to class members whose books and inserts have been digitized.

If a book, either in-print or out-of-print, was created as a work-for-hire, the Registry will remit 100\% of revenues to the publisher.\textsuperscript{261} The authors of works-for-hire will not actually be parties to the Settlement because they do not have a copyright interest in the work.\textsuperscript{262} Thus, for whatever it is worth, these authors are not bound by the terms of the Settlement.

The remaining provisions relating to the allocation of revenues depend on a book’s in-print/out-of-print status. For books in-print, the Registry will make all payments to the publisher and the publisher is then expected to pay the author based upon the royalty rate agreed upon in their contract or other agreement.\textsuperscript{263} Most authors who object to their allocation under this provision can initiate a dispute under Article VII of the APP, however, the authors of Educational Books—books primarily intended for sale to educational markets for use in educational programs—do not get the benefit of Article VII.\textsuperscript{264}

Assuming that Article VII applies, once the Registry is notified of a dispute, it will suspend payments until the matter is resolved.\textsuperscript{265} Either party may submit the dispute to the Registry for arbitration under Section 9.3 of the Settlement if no amicable resolution can be achieved.\textsuperscript{266}

\textsuperscript{259} APP § 6.1(c).
\textsuperscript{260} APP § 5.5.
\textsuperscript{261} APP § 6.2(b).
\textsuperscript{262} § 1.142 (defining “Settlement Class”).
\textsuperscript{263} APP § 5.5.
\textsuperscript{264} Id. See § 1.3 (defining Educational Books for the purposes of the APP).
\textsuperscript{265} APP § 7.2(a).
\textsuperscript{266} APP § 7.2(b).
arbitration, the arbitrator will look first to the terms of the written contract, and will only look to extrinsic evidence if the contract does not contain conclusive language.\(^{267}\) While these provisions may be consistent with many author-publisher contracts, they are clearly inconsistent with agreements which do contain merger clauses or provide for mandatory arbitration. Some authors may see these procedures as overly restrictive or favorable to publishers.

The allocation of revenue for out-of-print books does not appear to rest on the terms of the relevant author-publisher contract. Presumably because a large number of author-published contracts would contain no such provision. If the relevant rights have reverted to the author, or the book is deemed to be author-controlled, the Registry will remit 100% of revenues to the author.\(^{268}\) Otherwise, if an out-of-print book was first published prior to 1987, the Registry will pay 65% of revenues to the author and 35% to the publisher.\(^{269}\) Revenues are to be divided equally between authors and publishers of out-of-print books published during or after 1987.\(^{270}\) There is no indication in the agreement as to why the year 1987 was chosen for this purpose. Both authors and publishers may take issue with this fixed payment schedule. These fixed percentages may be at odds with prior author-publisher arrangements or understandings regarding royalties, and may be grossly out of proportion as to the amounts agreed upon in the publishing agreement.

The assumption up to this point has been that all of the relevant rightsholders are registered with the Registry. However, in all likelihood a substantial number of authors and publishers will fail to register at least some of their titles. Potential rightsholders must register their interest in

\(^{267}\) APP § 7.2(c). Section 7.3 also defines the fee structure for disputes; giving power to the Registry to set filing fees in order to avoid financially burdening the Registry. APP § 7.3. For disputes between Authors and Publishers, the party submitting the dispute to the arbitrator must pay a filing fee of $300, $150 of which they will get back if they prevail. APP § 7.3(a). The Registry will set fees for disputes among non-Publisher Rightsholders, and all other costs will be borne by the Registry. APP § 7.3(b)-(c). Note also that if an In-Print book was published under an author-publisher contract executed prior to 1992 that has not been amended to expressly address electronic rights, the arbitrator’s ruling as to the allocation of revenues for electronic rights will also apply to other uses of that book in any other Google programs that are the same as or substantially identical to that Display Use. APP § 10.1

\(^{268}\) APP § 5.1, APP § 6.2(a).

\(^{269}\) APP § 6.2(c)(i). The publication year is determined by the earliest copyright year that appears in the book, or if none is listed, then the first year of publication of the book. APP § 6.2(c)(iii).

\(^{270}\) APP § 6.2(c)(ii).
In order to receive any payments under the Settlement. The mechanics of registration are not burdensome for any particular book, but they might prove challenging for publishers with poor record-keeping and extensive back-catalogues.

In order to register, claimants must provide their contact information, social security number or equivalent, and an attestation of their good faith belief of ownership of a copyright interest in the book or the insert that is the subject of the claim. To claim an interest in relation to a book, a claimant must: identify whether she is the author, publisher, heir, successor or assignee, or other type of rightsholder of the book; whether such book is a work-for-hire. A claimant must also provide information sufficient to identify the book, including, either the title of the book and the name(s) of the author(s) and/or editors or the book’s ISBN. Claimants must also state their belief that either the book is not a “United States work,” as defined in 17 U.S.C. § 101 or that the book was registered with the U.S. Copyright Office by the Notice Commencement Date. In addition, publisher claimants must state whether they are “highly confident” based on the individual book contract that the rights to the book have not reverted to the author or whether they are “confident”, based on the type of book or type of contract for the book, that the rights to the book have not reverted. In contrast, authors and their successors in title are merely required to state either that the book has or has not reverted or that they do not know whether such the book has reverted.

The timing of a claimant’s registration is consequential. Although rightsholders may register to participate in the Settlement at any time, to be eligible for a cash payment pursuant to Section 5.1 of the Agreement, which provides for cash payment to class members whose works have been digitized, a rightsholder must register within one year of the Notice Commencement Date. Furthermore, rightsholders effectively abandon any revenues and attributable to their works which are unclaimed after five years. The unclaimed funds provision of the agreement provides that revenues, other than subscription revenues, which are not claimed within

271 § 13.1.
272 § 13.1(a).
273 § 13.1(c).
274 Id.
275 Id.
276 Id.
277 §13.3.
five years will be distributed by the Registry as follows:278 (1) to the Registry itself to defray reasonable and necessary operational expenses; (2) on a proportional basis to the registered rightsholders in order offset the 10% deduction for Google’s operating costs built into the net advertising revenues and net purchasing revenues;279 (3) to not-for-profit entities chosen by the Registry after consultation with Google and the participating and cooperating libraries.280 Funds attributable to the subscription revenue models will be distributed by the Registry in accordance with the Plan of Allocation if they are remain unclaimed for at least five years.281 The Plan of Allocation provides for the reallocated of such funds to registered rightsholders to pay the target inclusion fees—$200 per book for inclusion in the institutional subscription—and then to defray reasonable and necessary operational expenses of the Registry.282 The APP adds to the unclaimed funds provision in Section 6.3(a) of the Settlement. It stipulates that the unclaimed funds for out-of-print books attributable to authors should be distributed to registered rightsholders in the author subclass and that unclaimed funds attributable to publishers should be distributed to the publisher subclass.283 Thus in true Dickensian fashion, the Registry, registered rightsholders and certain selected charities all stand to benefit from the continued exploitation of orphans. One might therefore question the strength of these party’s incentives to diligently restore orphaned works to their true owners.

5. Some Implications of the APP

In many respects, the APP simply streamlines the incorporation of existing contract terms between authors and publishers into the framework of the Google Book Settlement. This is essential to the coordinating function of the Settlement and day to day operation of the Registry. The provisions of the APP are unlikely to be controversial in situations where rights have clearly reverted back to the original author, or where the original author has no rights at all because the work was a work made-for-hire. However, where the existing author-publisher contract arguably gives both parties some control over the exploitation of the work, or simply fails to make any provision for electronic rights, the APP effectively amends that contract.

278 Non-Subscription Revenues include Consumer Purchases, Advertising Revenue, Printing, and, if agreed, Print on Demand, Custom Publishing and PDF Download, respectively. APP § 6.3.
279 § 6.3. See also § 4.5(a)(i)-(ii).
280 § 6.3.
281 § 6.3.
282 1.1(e) of the Plan of Allocation, Attachment C to the Settlement Agreement.
283 APP § 8.1, APP § 8.2.
Whether the APP fills a contractual void or overwrites provisions on an existing contract it will inevitably disadvantage some authors and some publishers. These new terms do not appear to systematically disadvantage either authors or publishers, but there may well be individual authors or publishers who would strenuously object to them. The public may also be disadvantaged by the conservative nature of the APP. In a range of scenarios in which both authors and publishers maintain an interest in the work and have conflicting preferences, the APP takes a lowest common denominator approach and provides for the most restrictive access at the highest price. In the aggregate, these disadvantages are probably offset by the greater efficiency and increased public access facilitated by the APP.

PART III. CONCLUSION—EVALUATING THE SETTLEMENT IN LIGHT OF THE FAIR USE COUNTERFACTUAL

As discussed in Part I, the most likely litigation outcome was that the digitization of library books for inclusion in a book search database would constitute a fair use so long as the final product—the search results displayed to end-users—was not infringing; and so long as copyright owners were provided with a low-cost mechanism to withdraw their books from the database. It must be stressed that this was just one possible outcome, and things certainly could have gone much worse from Google’s point of view. Nonetheless, this counterfactual gives us a useful basis upon which to evaluate the Settlement agreement.

To make this comparison properly requires an analysis of more than just the initial rights and freedoms created by such a fair use ruling. It also requires some investigation of how Google and copyright owners would have been likely to have negotiated in the shadow of such a ruling. In other words, there is a difference between the ‘fair use ruling’ and the subsequent moves the players would make in a ‘fair use world’.

Under the hypothetical fair use ruling that forms the basis of the counterfactual, Google would have the right to continue to make non-expressive uses of printed library books by digitizing and analyzing these works. It would also have the right to continue to report such results in relation to these books provided that the search results themselves did not amount to a substantial copy of any particular work. The fair use ruling would not have allowed any substantial previews, with the exception of works in the public domain which are of course unrestricted. In addition,
the hypothetical fair use ruling would not, by itself, provide a single cent of revenue for any copyright owner.

Now consider the fair use world and how it compares to the Settlement. In a fair use world, Google’s basic digitization, analysis and reporting would continue, except that some copyright owners would presumably opt-out of the database altogether. This is essentially the same as the class-action Settlement. In a fair use world, the display of brief snippets of text as part of a menu of search results would continue to be non-infringing provided that it did not cross over the threshold of substantial similarity. The Settlement agreement is very similar in this regard except that rather than relying on the somewhat imprecise copyright law concept of substantial similarity, the length of the snippets allowed is narrowly defined under the Settlement. Under both the Settlement and in a fair use world, some copyright owners would probably opt-out of snippets.

It is likely that, in a fair use world, some copyright owners would have expressly allowed Google to make more substantial uses of their works, including the display of substantial previews and possibly even if the kinds of activities envisaged under the “New Revenue Models” discussed in the Settlement. However, the number of copyright owners who would have taken such action remains uncertain. A key difference between the Settlement and the fair use world hypothesized here is that the Settlement makes it much easier for rightsholders to opt in to expansive uses or their works. Indeed, many of the defaults established by the Settlement do not even require an opt-in, rather they place the burden on the rightsholder to opt-out. The Settlement sets a default position which allows significant previews of out-of-print/not-commercially-available books. It also subjects most of these works to consumer purchases and institutional subscriptions—uses that potentially substitute for existing paper collections rather than merely complimenting them.

The Settlement also contains a number of mechanisms through which authors, and the Registry acting on behalf of authors, can increase the extent of the previews they allow. It also provides a platform for negotiation for the “New Revenue Models”. Beyond the specific extensions provided for in the agreement, one of the key differences between the fair use world and the Settlement is the institutional framework erected by the agreement. The Settlement establishes contractual mechanisms for rightsholders to specify their preferences and creates an important institution to facilitate rightsholder decision making and collective negotiation. This institutional framework allows for all of the same transactions that might occur in a fair
use world, but at a significantly lower cost of transacting. The expectation in a fair use world is that some authors to profit from opting into more extensive uses, or from promising not to opt-out of the basic uses. Nonetheless, the opportunities for copyright owner remuneration under the Settlement appear to be much more significant.

The other key institutional aspect of the Settlement is the Author-Publisher Procedures. Although these procedures allocate control and beneficial ownership between authors and publishers, they exist primarily for the benefit of the Registry and Google. The APP provides a mechanism for incorporating existing author-publisher contractual terms into the Google administrative framework. It also specifies new contractual terms which fill critical gaps and ambiguities in a number of legacy publishing agreements. The APP provides a set of mechanisms through which the Registry can reconcile potentially conflicting directions regarding the electronic exploitation of a registered work. It also provides the Registry with a formula for allocating revenues between authors and publishers in situations where many existing contracts make no adequate provision.

What does this counterfactual comparison imply for an evaluation of the Settlement? Comparing the Settlement to the most likely outcome of the author and publisher lawsuits gives us an important baseline from which to evaluate the merits of the Settlement agreement. Much of the analysis in this Article suggests that the Settlement negotiated between the parties is not far removed from the most likely resolution a court would have adopted. In other respects, however, the Settlement and the fair use counterfactual are worlds apart.

The carefully negotiated settlement provides a much more elaborate institutional framework event would have emerged following a fair use ruling in Google’s favor. It almost certainly allows for more efficient exploitation of printed library books. In addition, whereas a fair use ruling would have effectively unblocked access to existing works for the limited purpose of digitization, indexing and search; the Settlement provides more access to existing works through electronic consumer purchases, the institutional subscription service and other products and services yet to be developed. On the other hand, whereas a fair use precedent would have cleared a significant legal obstacle for all potential entrants to the book search field, the Settlement only licenses Google.

At this point in time we simply lack the empirical data required to determine whether narrower rights available to all would be preferable to
broader rights concentrated in the hands of one commercial entity. My own intuitive assessment of the Settlement is that the gains more than outweigh the losses. Still, I cannot help but wonder what might have been, both for the development of book search and the development of fair use jurisprudence at large.