

An Empirical Study of U.S. Copyright Fair Use Opinions, 1978-2005

Barton Beebe[†]

Abstract

Section 107 of the Copyright Act of 1976 establishes the affirmative defense to copyright infringement of “fair use,” by far the most enigmatic doctrine in U.S. copyright law and by far the most important. Without it, much of our economic and communicative action would constitute copyright infringement. Yet despite the importance of the fair use defense, and despite the enormous amount of scholarly attention that it has received, we continue to lack any systematic, comprehensive account of our fair use case law and the actual state of our fair use doctrine. Instead, our conventional wisdom derives from a small set of conventionally agreed upon leading cases. This Article presents the results of the first empirical study of our fair use case law to show that much of our conventional wisdom about that case law is wrong. Working from a data set consisting of all reported federal opinions that made substantial use of the Section 107 four-factor test for fair use through 2005, the Article shows which factors and subfactors actually drive the outcome of the fair use test in practice, how the fair use factors interact, how courts inflect certain individual factors, and the extent to which judges stampede the factor outcomes to conform to the overall test outcome. It also presents empirical evidence of the extent to which lower courts either deliberately ignored or were ignorant of the doctrine of the leading cases, particularly those from the Supreme Court. Based on these descriptive findings, the Article prescribes a set of doctrinal practices that will improve courts’ adjudication of the fair use defense.

Word count, including abstract, footnotes, tables, and appendix: 27,722

TABLE OF CONTENTS

INTRODUCTION.....	2
I. BACKGROUND.....	7
A. Section 107 of the Copyright Act of 1976	7

B. Courts' Mechanical Application of the Four-Factor Test	10
II. SUMMARY STATISTICS	13
A. Distributions of the Opinions.....	13
1. Distribution of the Opinions by Year and Venue	14
2. Distribution of District Court Opinions by Posture.....	17
3. Distribution of the Opinions by Subject Matter	20
B. Reversal, Dissent, and Appeal Rates.....	21
C. Fair Use Win Rates	22
III. INTERFACTOR ANALYSIS.....	28
A. Correlation and Regression Analysis.....	28
B. Word Count Analysis.....	32
C. Stampeding	34
IV. INTRAFACITOR ANALYSIS.....	38
A. Factor One: Purpose and Character of the Use	40
1. The Commerciality Inquiry	41
2. The Transformativeness Inquiry.....	45
3. The Bad Faith Inquiry	48
4. The Preambular Purposes Inquiry	49
B. Factor Two: Nature of the Copyrighted Work.....	50
1. The Creative/Factual Work Inquiry.....	51
2. The Published/Unpublished Work Inquiry.....	52
C. Factor Three: Amount and Substantiality of the Use	54
D. Factor Four: Effect on the Market.....	55
V. CONCLUSION.....	59
APPENDIX: THE COLLECTION AND CODING OF THE OPINIONS	61

INTRODUCTION

Section 107 of the Copyright Act of 1976¹ establishes the affirmative defense to copyright infringement of “fair use.”² This affirmative defense represents the most important—and amorphous—limitation on the otherwise extraordinarily broad rights granted to copyright owners under Section 106 of

[†]Associate Professor of Law, Benjamin N. Cardozo School of Law, Yeshiva University. B.A., University of Chicago; Ph.D., Princeton University; J.D., Yale University. Thanks to Michael Carroll, Thomas Cotter, Rochelle Dreyfuss, Jim Gibson, Wendy Gordon, Justin Hughes, Robert Kasunic, Michael Madison, Matthew Sag, and Jessica Silbey for comments on and help with the project. This project greatly benefited from presentations to the London School of Economics and Political Science, the Boalt Hall Intellectual Property Scholarship Colloquium, and the Intellectual Property Scholars Conference 2006. Thanks also to Deepica Kapoor and Carl Regelmann for help in collecting the cases. I used Intercooled Stata 9.2 for Windows to conduct the quantitative data analysis presented in this Article and Atlas.ti 5.2.9 to conduct the qualitative data analysis, including word count analysis. Comments welcome: barton@bartonbeebe.com.

¹ Pub.L. No. 94-553, 90 Stat. 2541 (codified as amended at 17 U.S.C. § 101 (2000)).

² See 17 U.S.C. § 107.

the Act.³ Much of our economic and communicative action would constitute infringement of those rights but for the affirmative defense of fair use. It is the ill-defined and unpredictable point of flexibility, the “dangerous supplement,”⁴ that makes the rest of our copyright law possible—and bearable.⁵ Section 107 offers no definition of “fair use,” and none is offered elsewhere in the Act, or indeed in the case law.⁶ The section does, however, enumerate four broadly-worded factors that courts “shall” consider in determining whether a use is “fair” and thus non-infringing:

- (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and (4) the effect of the use upon the potential market for or value of the copyrighted work.⁷

For better or worse, these factors form the core of our fair use doctrine and functionally define what fair use is. For all of the ambiguity of their statutory language, they are what stand between us and what some have called, with only a touch of hyperbole, the “tyranny of copyright.”⁸

Section 107 and the concept of fair use more generally have attracted an enormous amount of scholarly attention, ranging from broad social-theoretical treatments⁹ to precise analyses of the leading cases and their impact.¹⁰ Nearly all of this commentary has been highly critical of Section 107’s four-factor test

³ See 17 U.S.C. § 106.

⁴ See JACQUES DERRIDA, *OF GRAMMATOLOGY* 141-64 (Gayatri Chakravorty Spivak trans. 1976) (describing a “dangerous supplement” as, among other things, an area of deconstructive ambiguity and ambivalence in a theoretical edifice that threatens always to undermine the foundations of that edifice).

⁵ Cf. R. Polk Wagner, *The Perfect Storm: Intellectual Property and Public Values*, 74 FORDHAM L. REV. 423 (2005) (discussing how uncertainty affects fair use litigation and doctrine).

⁶ Cf. *Rosemont Enterprises, Inc. v. Random House, Inc.*, 366 F.2d 303, 306 (2d Cir. 1966) (“Although no definition of fair use that is workable in every case has ever evolved, a frequently quoted definition of fair use is a privilege in others than the owner of a copyright to use the copyrighted material in a reasonable manner without his consent, notwithstanding the monopoly granted to the owner (by the copyright).” (citations omitted)).

⁷ *Id.*

⁸ See, e.g., Robert S. Boynton, *The Tyranny of Copyright?*, N.Y. TIMES, Jan. 25, 2004, § 6 (Magazine), at 40. See also Yochai Benkler, *Free as the Air to Common Use: First Amendment Constraints on Enclosure of the Public Domain*, 74 N.Y.U. L. REV. 354 (1999); James Boyle, *The Second Enclosure Movement and the Construction of the Public Domain*, 66 LAW & CONTEMP. PROBS. 33 (2003).

⁹ See, e.g., William W. Fisher III, *Reconstructing the Fair Use Doctrine*, 101 HARV. L. REV. 1659 (1988); Wendy J. Gordon, *Fair Use as Market Failure: A Structural and Economic Analysis of the Betamax Case and Its Predecessors*, 82 COLUM. L. REV. 1600 (1982); Michael J. Madison, *A Pattern-Oriented Approach to Fair Use*, 45 WM. & MARY L. REV. 1525 (2004).

¹⁰ See, e.g., Jessica Litman, *The Sony Paradox*, 55 CASE W. RES. L. REV. 917 (2005); Frank Pasquale, *Breaking the Vicious Circularity: Sony’s Contribution to the Fair Use Doctrine*, 55 CASE W. RES. L. REV. 777 (2005); Pamela Samuelson, *The Generativity of Sony v. Universal: The Intellectual Property Legacy of Justice Stevens*, 74 FORDHAM L. REV. 101 (2006); Peter S. Menell & David Nimmer, *Unwinding Sony* (working paper on file with author). See also *infra* note 63 discussing the annual ratio of law review articles addressing fair use to actual federal court cases doing the same during the period sampled.

and how courts have applied it.¹¹ Yet, remarkably, we continue to lack any systematic, comprehensive account of our fair use case law.¹² Instead, like the “great men”¹³ approach to history, we pursue a leading cases (or “usual suspects”) approach to fair use. This anecdotal method, one essentially of connoisseurship, derives conventional wisdom about our case law from a limited aristocracy of hand-picked opinions appearing primarily in the *U.S.* and *Federal Reports*—or in the student casebooks.¹⁴ Whether these opinions have any influence on or are representative of the true state of our fair use doctrine—the actual rather than the theoretical state of it, the doctrine as it is practiced in the courts—remains an open, and strangely unasked, question.

To answer this and other questions, this Article presents the results of the first empirical study of our fair use case law.¹⁵ Working from a data set

¹¹ See, e.g., WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF INTELLECTUAL PROPERTY LAW 115 (2003) (criticizing the four-factor approach); Madison, *supra* note 9, at 1564 (“[T]he facial emptiness of the statutory language means that alone, it is almost entirely useless analytically, except to the extent that it structures the collection of evidence that a court might think relevant to its decision.”); *id.* at 1586 (“Despite the small sample of cases reviewed in the last Part, the point should be clear that across a range of fair use cases, the Supreme Court’s formal jurisprudence has encouraged the courts of appeals, and presumably the district courts following their lead, to abstract the fair use inquiry to the point of incoherence.”); Matthew Sag, *God in the Machine: A New Structural Analysis of Copyright’s Fair Use Doctrine*, 11 MICH. TELECOMM. TECH. L. REV. 381, 434 (2005) (referring to Section 107 as a “vague and open-ended standard”). See also Michael Carroll, *Fixing Fair Use*, 85 N.C. L. REV. __, __ (2007) (forthcoming) (“[L]eading courts and commentators generally acknowledge that the four-factor test as interpreted provides very little guidance for predicting whether a particular use will be deemed fair.”). But see William F. Patry & Shira Perlmutter, *Fair Use Misconstrued: Profit, Presumptions, and Parody*, 11 CARDOZO ARTS & ENT. L.J. 667, 667 (1992) (“The thorniness of fair use is not due to any failure on the part of the legislature or the judiciary. To the contrary, it is inherent in the common law nature of the doctrine as a multifaceted process by which courts can finely calibrate not only the equities between the parties in a given case, but also the fundamental public policies at stake in drawing the line between private property and free use.”).

¹² Cf. Rochelle Dreyfuss, *Games Economists Play*, 53 VAND. L. REV. 1821, 1825 (2000) (“Phenomena that generate data, like collective rights organizations, are amply studied; amorphous doctrines, such as fair use, which have enormous legal significance but little by way of hard facts, do not receive as much attention from pure economists.”).

¹³ See, e.g., THOMAS CARLYLE, ON HEROES, HERO-WORSHIP AND THE HEROIC IN HISTORY 29 (Carl Niemeyer ed., Univ. Neb. Press 1966) (1843) (“The history of the world is but the biography of great men.”).

¹⁴ We currently have one study of a cross-sectional sample of fair use opinions. See David Nimmer, “*Fairest of Them All*” and *Other Fairy Tales of Fair Use*, 66 Law & Contemp. Probs. 263, 281 (2003). Nimmer studied sixty fair use opinions published in the *Federal Supplement* or the *Federal Reports* between the March 7, 1994 opinion date of *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569 (1994), and the February 6, 2002 opinion date of *Kelly v Arriba Soft Corp.*, 280 F.3d 934 (9th Cir. 2002). As Nimmer admits, his method was idiosyncratic. He read the facts stated in the opinions and then, based on those facts, decided for himself which party he thought each factor should favor. He then compared his own factor-specific findings to the overall finding of the court on the fair use question. See *id.* at 267 n. 28. Based on a data set assembled and provided to me by Matthew Sag, I conducted a logistic regression of the court’s determination of the outcome on Nimmer’s determinations of the outcomes of the four factors. On this regression, the only significant factor outcome was the second, going to the nature of the plaintiff’s work, and the coefficient was negative. Thanks to Matthew Sag for informing me of this curious result.

¹⁵ The method employed here is generally referred to as “content analysis.” See Mark A. Hall & Ronald F. Wright, *Systematic Content Analysis of Judicial Opinions* (working paper on file with author). See also Fred Kort, *Content Analysis of Judicial Opinions and Rules of Law*, in *JUDICIAL DECISION-MAKING* (Glendon Schubert, ed. 1963); Reed C. Lawlor, *Fact Content Analysis of Judicial Opinions*, 8 JURIMETRICS J. 107 (1966-1968).

consisting of all reported federal opinions that made substantial use of the Section 107 four-factor test from the January 1, 1978 effective date of the Copyright Act through 2005, the Article shows that much of our conventional wisdom about our fair use case law, deducted as it has been from the leading cases, is wrong. The dual purpose of the Article is systematically to induce from the population of Section 107 opinions what our fair use doctrine actually is in practice, and critically to induce from that population of opinions what it ought to be.

The Article proceeds as follows. Part I provides background on Section 107 and evidences in the data the curious origins of the highly routinized fashion in which judges typically apply its four factors. Part II sets forth summary statistics on the 306 opinions studied, including their distribution along the dimensions of time, venue, posture, and kinds of copyrightable works at issue. The Part reveals the full extent to which the Second and Ninth Circuits dominate our fair use case law. It also identifies and seeks to explain the remarkable increase in the prevalence of fair use summary judgment opinions that began in the mid-1990s and has continued to the present, and the just as remarkable absence of fair use bench trial opinions during the same period. The Part then reports the reversal, dissent, and appeal rates in the fair use case law and finds that notwithstanding strong conventional wisdom to the contrary, none of these is especially high as compared to other areas of law. Finally, and perhaps most interestingly, the Part sets out the fair use win rates (i.e., the proportion of opinions that found fair use) in the case law. For certain postures, these rates were exceptionally low. The Part seeks to explain why.

Parts III and IV then focus on how judges used the four factor test. Part III engages in an *interfactor* analysis of the test as applied. It focuses on the factor outcomes and shows how these outcomes interacted with each other and with the overall test outcome. Correlation and regression analysis reveal the surprising extent to which the outcomes of the first and especially the fourth factors appear to drive the outcome of the test. The Part also reports the results of word-count analysis of the opinions across time. The Part concludes by considering the extent to which judges stamped¹⁶ the factor outcomes to

¹⁶ The metaphor of stampeding is adopted from Barton Beebe, *An Empirical Study of the Multifactor Tests for Trademark Infringement*, 95 CAL. L. REV. 1581 (2006). The hypothesis is that judges decide the outcome of a multifactor test based on a limited number of core factors, possibly only one; the judge then tends to insure that the most if not all of the remaining factors follow the lead of this dispositive factor. Alternatively, the hypothesis is that certain multifactor tests are by their own nature prone to stampeding, primarily because the factors are redundant of each other. The idea of stampeding is based in part on the “coherence-based reasoning model” proposed by Dan Simon and others, see, e.g., Dan Simon, *A Third View of the Black Box: Cognitive Coherence in Legal Decision Making*, 71 U. CHI. L. REV. 511 (2004), and in part on the fast-and-frugal heuristics research of Gerd Gigerenzer and others, see, e.g., Gerd Gigerenzer & Daniel G. Goldstein, *Reasoning the Fast and Frugal Way: Models of Bounded Rationality*, 103 PSYCH. REV. 650 (1996). Cf. Nimmer, *supra* note 14, at 281 n. 62 (“Alternatively, as courts work their way through the four factors, at some point they decide what the ultimate conclusion should be—which, in turn, molds the way that they reach resolution as to which direction each factor points.”).

conform to the overall test outcome. The data show that, again, notwithstanding strong conventional wisdom to the contrary, and notwithstanding the example of certain leading cases, judges do not generally stampede the factor outcomes. Rather than make a fair use determination first and then “align the four factors to fit that result as best they can,”¹⁷ courts appeared quite willing to call the factor outcomes as they saw them, even when those outcomes did not support the overall test outcome. This is a highly encouraging finding.

Part IV then turns to an *intrafactor* analysis of the fair use test as applied. It looks inside each of the four factors to determine what subfactor considerations animated courts’ determination of the factor outcomes, and through those factor outcomes, the outcome of the overall fair use test. Regression analysis as well as simple descriptive statistics show that certain subfactor findings are far more important to the outcome of the fair use test than is generally thought, and that others thought to be important actually had no significant influence on the outcome of the test. For example, a finding that the plaintiff’s work was factual in nature or published appeared strongly to influence a court to find fair use, as did a finding that the defendant’s use was for a non-commercial purpose. Meanwhile, the data suggest that a finding that the defendant’s use was for a commercial (rather than non-commercial) purpose had no significant influence on the outcome of the test, and a finding that the defendant used the entirety of the plaintiff’s work was far from dispositive. More generally, and more disturbingly, the Part shows the extent to which lower courts repeatedly either deliberately ignored or were ignorant of basic, albeit dictic, Supreme Court precedent. Indeed, the data show that the lower courts repeatedly and systematically inverted Supreme Court dicta to favor the defendant, so that if the Court stated, for example, only that “not X” favors the plaintiff, the primary lesson the lower courts would draw from this is that “X” favors the defendant. This may come as a pleasant surprise to fair use maximalists, who are generally pessimistic about the state of our fair use case law and the judges who produce it. But when the lower courts systematically resist the authority of the Supreme Court, and on issues of central importance to an area of law, this is cause for concern, regardless of the policy outcome. The Part suggests that the indiscipline of the lower courts is largely the fault of the Supreme Court itself and its repeated unwillingness explicitly to correct its own past mistakes in its fair use opinions. This has left the lower courts with an array of “accumulating precedent”¹⁸ from which to choose.

Part V concludes by reflecting on the prescriptive implications of the study’s findings. In the mass of everyday fair use case law, and in the

¹⁷ Nimmer, *supra* note 14, at 281.

¹⁸ See generally Stefanie Lindquist & Frank B. Cross, *Empirically Testing Dworkin’s Chain Novel Theory: Studying the Path of Precedent*, 80 N.Y.U. L. Rev. 1156 (2005).

interstices of the four factors, we can identify what considerations most commonly formed the basis of judges' fair use determinations. This we can welcome as empirical precedent, that is, as precedent based on the empirical analysis of how the population of judges who have previously employed a legal doctrine did so. In the near term, we are stuck with Section 107, "botched job"¹⁹ that it may be, and the task is how to make the best of what we have. But in the long term, the more important task is to induce—both systematically and critically—from the Section 107 case law we have what form a new Section 107 might someday take.

I. BACKGROUND

As a foundation for the discussion that follows in subsequent Parts, this Part briefly reviews the language and legislative history of Section 107. It then describes the highly routinized analytical and rhetorical style that judges employed when they applied the Section 107 test. This style greatly facilitated reliable coding of the data set.²⁰

A. Section 107 of the Copyright Act of 1976

Structurally, the Copyright Act leans towards protection; it gives broadly and takes away narrowly. Sections 106 enumerates in expansive terms the exclusive rights that a copyright owner enjoys in a copyrighted work.²¹ Sections 107 through 122 then set forth "limitations" on these exclusive rights. Most of these limiting sections contain highly specific, even regulatory language.²² Section 107, however, among the briefest of the limiting sections,²³ is extraordinarily broad in scope. The Section states in full:

Notwithstanding the provisions of sections 106 and 106A, the fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use, the factors to be considered shall include—

¹⁹ Lloyd L. Weinreb, *Fair's Fair: A Comment on the Fair Use Doctrine*, 103 HARV. L. REV. 1137, 1151 (1990).

²⁰ For a discussion of the process of collecting and coding the opinions that comprised the data set for this study, see the Appendix.

²¹ 17 U.S.C. § 106.

²² See generally Joseph P. Liu, *Regulatory Copyright*, 83 N.C. L. REV. 87 (2004).

²³ The only limiting section shorter in word count is Section 120, involving the scope of exclusive rights in architectural works. See 17 U.S.C. § 120.

- (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
- (2) the nature of the copyrighted work;
- (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
- (4) the effect of the use upon the potential market for or value of the copyrighted work.

The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the above factors.²⁴

Section 107 thus consists of three parts:²⁵ a preamble that identifies “the fair use of a copyrighted work” as an exception to the copyright owner’s exclusive rights and that gives examples of fair “purposes,” a list of four factors that courts must consider in determining whether a use is a fair use,²⁶ and an additional statement regarding unpublished works that was added in 1992.²⁷

A detailed account of the legislative history of Section 107 is beyond the scope of this Article and is available elsewhere.²⁸ It is interesting to note, however, that early on in the two decades of negotiations that led to the Act, a variety of interested parties urged that the Act not reference fair use at all, or that if it did, its reference take the form of a “bare bones”²⁹ one sentence limitation. Their concern was that statutory recognition of fair use would “freeze” or otherwise unpredictably alter the judge-made doctrine.³⁰ Indeed, in 1965, after draft provisions similar in structure to the current Section 107 raised

²⁴ 17 U.S.C. § 107.

²⁵ Cf. Weinreb, *supra* note 19, at 1139 (“Congress adopted three considerably inconsistent ways of doing nothing: simple reference to fair use, specification of what is fair use by illustrative examples, and prescription of nonexclusive ‘factors to be considered’ in determining whether a particular use is fair. As Hercule Poirot observed about the murder on the Orient Express, the problem is not that there are too few clues but that there are too many.”).

²⁶ Judge Posner was quite dismissive of the factors in *Ty, Inc. v. Publications Intern. Ltd.*, 292 F.3d 512, 522 (7th Cir. 2002). On remand, District Court Judge Zagel nevertheless explained to the defendant, who adopted Judge Posner’s approach “to its detriment,” that “regardless of what PIL claims the Seventh Circuit did, statutory and case authority mandate that the four fair use factors must be considered and applied in every case, including this remand.” *Ty, Inc. v. Publications Intern., Ltd.*, 333 F.Supp.2d 705, 715-16 (N.D.Ill. 2004).

²⁷ Pub. L. No. 102- 492, 106 Stat. 3145 (1992). See generally Kenneth D. Crews, *Fair Use of Unpublished Works: Burdens of Proof and the Integrity of Copyright*, 31 ARIZ. ST. L.J. 1, 42-66 (1999) (discussing the legislative history of the 1992 amendment and its impact on the fair use case law).

²⁸ See WILLIAM F. PATRY, THE FAIR USE PRIVILEGE IN COPYRIGHT LAW 261-365 (2d ed. 1995). On the legislative history of the Copyright Act in general, see Jessica Litman, *Copyright Legislation and Technological Change*, 68 OR. L. REV. 275 (1989); Jessica D. Litman, *Copyright, Compromise, and Legislative History*, 72 CORNELL L. REV. 857 (1987).

²⁹ See, e.g., *Copyright Law Revision: Hearings on S. 1006 Before the Subcomm. on Patents Trademarks and Copyrights of the Senate Comm. on the Judiciary*, 89th Cong., 51-52 (1965) (statement of the American Bar Association); *id.* at 1216 (statement of Association of American University Presses).

³⁰ See, e.g., DISCUSSION AND COMMENTS OF REPORT OF THE REGISTER OF COPYRIGHTS ON THE GENERAL REVISION OF THE U.S. COPYRIGHT LAW, COPYRIGHT LAW REVISION PART 3 51-52 (Comm. Print 1963) (hereinafter COPYRIGHT LAW REVISION PART 3).

a storm of objections (many of them the same objections we hear today), the negotiations focused on the following draft provision: “Notwithstanding the provisions of section 106, the fair use of a copyrighted work is not an infringement of copyright.”³¹ This proposed language failed to win consensus, however, and the next year, negotiations shifted back to the consideration of a provision taking the structure of a preamble and four factors.³²

The language of the preamble and factors changed in some cases dramatically through the course of the negotiations. As for the preamble, which was explicitly referenced by 22% of the 306 opinions studied, an initial 1963 draft proposed simple language that did little other than identify fair use as a “privilege.”³³ Three identical 1964 bills then proposed a more involved “legitimate purpose” requirement.³⁴ Dropped in 1965, the preamble returned in 1966 in a form quite similar to the final version of Section 107.³⁵ All that was missing was the parenthetical “(including multiple copies for classroom use),” which was added at the eleventh hour in early 1976.³⁶ As this addition reflects, the main controversy surrounding the section’s preambulatory language, if not the fair use section as a whole, concerned an emergent technology of the time: xerography, i.e., photocopying.³⁷ The negotiating parties may have overestimated the future significance to fair use of photocopying technology, however. Only fourteen opinions from nine cases addressed facts involving photocopying,³⁸ and only five of these opinions explicitly referenced the language of Section 107’s preamble.

The language of Section 107’s factors was largely drawn from Justice William Story’s 1841 circuit court opinion in *Folsom v. Marsh*,³⁹ an opinion whose influence on American fair use case law up to the 1976 Act we have probably overestimated, or so the data suggest,⁴⁰ but whose influence since is

³¹ H.R. 4347 and S. 1006, 89th Cong. (1st Sess. 1965). See also PATRY, *supra* note 28, at 277-296.

³² See PATRY, *supra* note 28, at 296-319.

³³ See COPYRIGHT LAW REVISION PART 3, *supra* note 30, at 5.

³⁴ See H.R. 11947, H.R. 12354, and S. 3008, 88th Cong., (2d Sess. 1964), reprinted in COPYRIGHT LAW REVISION PART 4, 1964 REVISION BILL WITH DISCUSSIONS AND COMMENTS, 89th Cong., 1-32 (Comm. Print 1965).

³⁵ See H.R. 4347, 89th Cong., § 107 (2d Sess. 1966).

³⁶ See PATRY, *supra* note 28, at 351.

³⁷ See *id.*

³⁸ Of course, a more optimistic interpretation of these data is that the preamble of Section 107, together with various privately-drafted guidelines, clarified the issue to a degree that litigation over it was rarely necessary. On the history and of privately-draft fair use guidelines, see Kenneth D. Crews, *The Law of Fair Use and the Illusion Fair-Use Guidelines*, 62 OHIO ST. L.J. 599 (2001).

³⁹ 9 F. Cas. 342 (No. 4,901) (CC Mass. 1841). See generally R. Anthony Reese, *The Story of Folsom v. Marsh: Distinguishing Between Infringing and Legitimate Uses*, in INTELLECTUAL PROPERTY STORIES 259 (Jane C. Ginsburg ed., 2005).

⁴⁰ A total of forty-four federal cases cited to *Folsom* from 1841 through 1977. This citation count is based on a search for the string “Folsom v Marsh” in the Westlaw “allfeds-old” database and the Lexis “Federal Courts Cases Before 1945, Combined” database for federal case law before 1945 (the temporal limit of both of these databases), and in the Westlaw “allfeds” database for federal case law after 1944. The Lexis pre-1945 database yielded five cases not yielded by the Westlaw pre-1945 database. For a rough scale by which to assess *Folsom*’s citation count from 1841 to 1977, consider that a Westlaw search of the federal case law

quite clear.⁴¹ A 1963 draft stated that “the following factors, among others, shall be considered,” and then listed out the four factors nearly as we have them today.⁴² Bills from 1964 narrowed the introductory statement (“the factors to be considered shall include”) and added the phrase “market for or” to factor four.⁴³ Finally, in early 1976, primarily to allay the concerns of educators,⁴⁴ the Subcommittee of the House Committee on the Judiciary responsible for the Copyright Act added the following phrase to factor one: “including whether such use is of a commercial nature or is for nonprofit educational use.”⁴⁵ Despite the committee’s best efforts at the time to downplay its significance in their reports,⁴⁶ this phrase has since caused no end of trouble, as we will see in Part III.

B. Courts’ Mechanical Application of the Four-Factor Test

Had the drafters of the statutory language of Section 107 known how judges would end up actually applying that language, they may very well have fallen back on the 1965 one-sentence provision after all. In the opinions studied, courts often acknowledged that the four factor test should not be applied formulaically;⁴⁷ as one court put it, the test does not “constitute an

from the same period for cases including the terms “fair use” and “copyright” yielded 125 cases, of which fifteen cited to *Folsom*. Early-twentieth-century treatises also failed to emphasize *Folsom*. See, e.g., RICHARD ROGERS BOWKER, COPYRIGHT: ITS HISTORY AND ITS LAW 252 (Boston, 1912) (referencing *Folsom* once, for a proposition relating to piracy, in a thirteen page discussion of fair use); WILLIAM BRIGGS, THE LAW OF INTERNATIONAL COPYRIGHT (London, 1906) (no reference to *Folsom*); RICHARD C. DE WOLF, AN OUTLINE OF COPYRIGHT LAW 140-167 (1925) (discussing fair use and failing to reference *Folsom*).

⁴¹ Of the 306 opinions sampled for this study, 47 opinions (15.4%) cited to *Folsom*, with 26.1% of the appellate opinions doing so, and 42.9% of the Supreme Court opinions doing so. In forty of these forty-seven opinions, the judge either quoted directly from *Folsom*, quoted indirectly from *Folsom* by noting that he was quoting from an opinion that was itself quoting from *Folsom*, or otherwise discussed the facts in *Folsom*. One of these forty purported to quote directly from *Folsom*, though, pedantically speaking, it did so in error. See *Suntrust Bank v. Houghton Mifflin Co.*, 268 F.3d 1257, 1261 n. 9 (11th Cir. 2001) (“In *Folsom v. Marsh*, 9 F. Cas. 342 (C.C. Mass. 1841), Justice Story created the concept of ‘fair use.’”). Justice Story did not use the term “fair use” in *Folsom*. The term of art first appeared in reported federal case law in *Lawrence v. Dana*, 15 F. Cas. 26, 58 (C.C. Mass. 1869).

⁴² See COPYRIGHT LAW REVISION PART 3, *supra* note 30, at 1-36. In his earlier Tentative Draft Report, the Register of Copyrights phrased the factors slightly differently. See PATRY, *supra* note 28, at 264.

⁴³ See H.R. 11947, H.R. 12354, and S. 3008, 88th Cong., (2d Sess. 1964), reprinted in COPYRIGHT LAW REVISION PART 4, 1964 REVISION BILL WITH DISCUSSIONS AND COMMENTS, 89th Cong., 1-32 (Comm. Print 1965).

⁴⁴ See *supra* note 11, at 678 (1992).

⁴⁵ See PATRY, *supra* note 28, at 351.

⁴⁶ See, e.g., H.R. Rep. No. 1476, at 66 (94th Cong. 1976) (“This amendment is not intended to be interpreted as any sort of not-for-profit limitation on educational uses of copyrighted works. It is an express recognition that, as under the present law, the commercial or nonprofit character of an activity, while not conclusive with respect to fair use, can and should be weighed along with the other factors in fair use decisions.”).

⁴⁷ See, e.g., *Weissmann v. Freeman*, 868 F.2d 1313, 1323 (2d Cir. 1989) (“Analysis begins not by elevating the statutory guides into inflexible rules, but with a review of the underlying equities.”); *Financial Information, Inc. v. Moody’s Investors Service, Inc.*, 751 F.2d 501, 508 (2d Cir. 1984) (“The four factors which the fair use statute identifies as relevant to a determination of whether the doctrine applies . . . are equitable considerations to be assessed and weighed by the court; they are not simply hurdles over which an

algorithm that enables decisions to be ground out mechanically.⁴⁸ Yet the data show that after an initial period of flexibility, judges shifted in the late 1980s towards a rhetorically quite formal and explicit treatment of the Section 107 factors. Figure 1 shows the proportion of opinions over time in which the court adopted the rhetorical practice of explicitly stating which party each factor favored.⁴⁹ Overall, judges engaged in this practice in 59.5% of the 306 opinions. Some judges also concluded their Section 107 analysis with a summary explicitly reviewing the valence of each factor. Overall, judges did so in 32.7% of the opinions.⁵⁰

Interestingly, the shift towards a routinized application of the Section 107 test roughly coincided with the Supreme Court's May, 1985 issuance of its opinion in *Harper & Row v. Nation*,⁵¹ in which the Court found that *The Nation*'s unauthorized publication of excerpts from President Ford's soon-to-be published autobiography was not a fair use. In her opinion for the majority, Justice O'Connor was deliberate in her application of Section 107, but she did not engage in the kind of mechanical rhetoric that we see in most fair use opinions from 1987 to the present.⁵² Justice Brennan's dissent, however, was

accused infringer may leap to safety from liability. Rather than a sequence of four rigid tests, the fair use analysis consists of a "sensitive balancing of interests.").

⁴⁸ Chicago Bd. of Educ. v. Substance, Inc., 354 F.3d 624, 629 (7th Cir. 2003) (Posner, J.). See generally Patry & Perlmutter, *supra* note 11, at 685 ("Nor need the first factor simply be resolved as 'pro' or 'con' fair use, depending on which aspect of the use predominates, its profit-making purpose or its scholarly character. Before fair use was incorporated in the statute, courts did not treat the factors as a checklist, with each factor reduced to a plus or minus. The mere listing of four factors in the 1976 Act, however, has led some courts improperly to take this approach.").

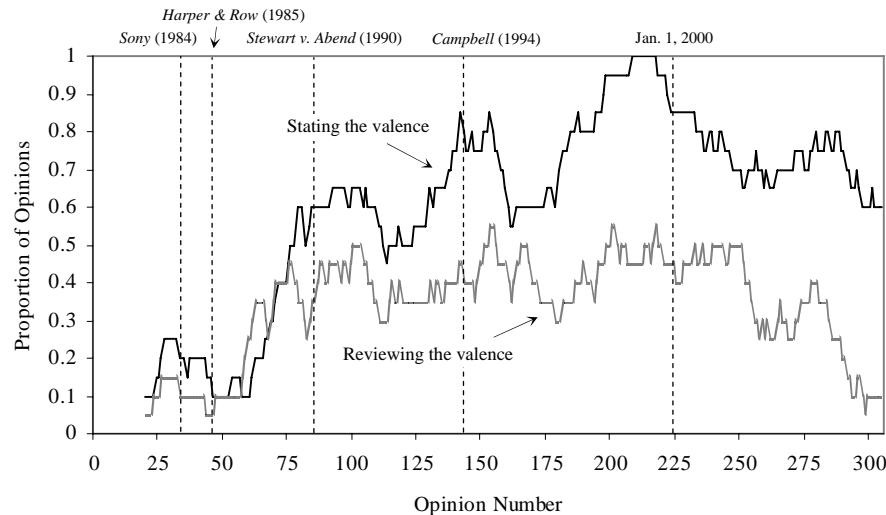
⁴⁹ See, e.g., Schiffer Pub., Ltd. v. Chronicle Books, LLC, No. 03-4962, 2004 WL 2583817, at *11-13 (E.D. Pa. 2004) ("Accordingly, the first fair use factor weighs in favor of Plaintiffs.... In sum, this Court cannot say that the second fair use factor militates strongly in either party's favor.... Consequently, the Court holds that the third fair use factor favors Plaintiff.... In sum, Defendants' use of Plaintiffs' work was not a fair use. Because the first, third, and fourth fair use factors favor Plaintiffs, and because none of the factors strongly favor Defendants, this Court holds that Defendants have not met their burden of proving that their appropriation of Plaintiffs' works constitutes fair use."); Batesville Services, Inc. v. Funeral Depot, Inc., No. 02-cv-01011, 2004 U.S. Dist. LEXIS 24336, at *22-26 (S.D. Ind. 2004) ("The first factor thus weighs heavily against the fair use defense.... This [second] factor is essentially neutral on Funeral Depot's fair use defense.... This [third] factor weighs against the fair use defense.... This [fourth] factor tends to weigh in favor of the fair use defense.").

⁵⁰ See, e.g., Kelly v. Arriba Soft Corp., 280 F.3d 934, 944 (9th Cir. 2002) ("Having considered the four fair use factors and found that two weigh in favor of Arriba, one is neutral, and one weighs slightly in favor of Kelly, we conclude that Arriba's use of Kelly's images as thumbnails in its search engine is a fair use."); Narell v. Freeman, 872 F.2d 907, 915 (9th Cir. 1989) ("In sum, the first factor weighs strongly in Narell's favor and the second factor slightly favors Freeman. However, a reasonable juror could only conclude that the third and final factors strongly favor Freeman."); Salinger v. Random House, Inc., 811 F.2d 90, 99-100 (2d Cir. 1987) ("On balance, the claim of fair use as to Salinger's unpublished letters fails. The second and third factors weigh heavily in Salinger's favor, and the fourth factor slightly so. Only the first factor favors Hamilton."); College Entrance Examination Bd. v. Cuomo, 788 F. Supp. 134, 143 (S.D.N.Y. 1992) ("In conclusion, given the fact that factor one favors the State, factor two favors GMAC, and factors three and four favor neither party, the court holds that GMAC has not demonstrated a likelihood of success on the merits of its copyright infringement claim.").

⁵¹ 471 U.S. 539 (1985).

⁵² See id. at 549-569.

FIGURE 1
20-OPIINION MOVING AVERAGE OF THE PROPORTION OF OPINIONS EXPLICITLY STATING THE VALENCE
OF THE FACTORS AND EXPLICITLY REVIEWING THE VALENCE OF THE FACTORS



highly mechanical in form.⁵³ It both explicitly stated the valences of each factor and explicitly summed up those valences in its conclusion.⁵⁴ This dissent arguably set American courts on a course towards a rhetorically more mechanical treatment of the Section 107 inquiry⁵⁵—notwithstanding the fact that the dissent itself urged courts to recognize that the Section 107 factors cannot “mechanistically resolve fair use issues.”⁵⁶

As further evidence of the generally routinized manner in which they employed the Section 107 test, judges rarely explicitly considered factors beyond the four listed in Section 107 and, with the exception of the second factor, rarely failed to consider less than all four factors.⁵⁷ Commentators have long called upon courts to look to additional factors,⁵⁸ such as “fairness,”⁵⁹ and in the opinions studied, the courts themselves frequently explicitly

⁵³ See id. at 619-635 (Brennan, J., dissenting).

⁵⁴ See id. at 634-35.

⁵⁵ Cf. Litman, *supra* note 28, at 862 n. 37 (noting that in *Sony* and *Harper & Row*, “the Supreme Court adopted a test that rigidifies the statute’s major flexibility principle.”).

⁵⁶ *Id.* at 624.

⁵⁷ Of the 306 opinions, 6.9% failed to consider the first factor, 17.7% failed to consider the second factor, 13.4% failed to consider the third factor, and 8.8% failed to consider the fourth factor.

⁵⁸ See, e.g., Gordon, *supra* note 9, at 1614 (calling upon courts to consider whether “(1) market failure is present; (2) transfer of the use to defendant is socially desirable; and (3) an award of fair use would not cause substantial injury to the incentives of the plaintiff copyright owner.”); Madison, *supra* note 9 (calling upon judges to conduct a “pattern-oriented” analysis of the fair use issue). See generally Fisher, *supra* note 9, at 1678-86 (discussing the relevance of various subfactor concerns to the fair use determination); Jennifer E. Rothman, *Customary Intellectual Property* (working paper on file with author) (discussing the relevance of industry “custom” to the fair use determination).

⁵⁹ See Weinreb, *supra* note 19, at 1140.

acknowledged that the Section 107 test was illustrative rather than limitative.⁶⁰ Yet only 17.0% of the opinions explicitly considered one or more additional factors, and only 8.8% stated that the additional was relevant to the fair use determination.

Thus, the methodical fashion in which judges used the four factor test to make their fair use determinations provides an orderly framework for systematic study. Whether this particular form of “mechanical jurisprudence”⁶¹ has been good for the law is, of course, another matter, one which I will address in later Parts. First, however, to build a foundation for a close analysis of the operation of the four factor test, I discuss in the next Part the general characteristics of the opinions.

II. SUMMARY STATISTICS

A. Distributions of the Opinions

In the twenty-eight years from the January 1, 1978 effective date of the 1976 Act to the conclusion of 2005, the federal courts produced 306 reported opinions from 215 cases that made substantial use⁶² of the Section 107 four-factor test. This averages out to 10.9 opinions per year during the twenty-eight year period, with an average of 4.6 opinions per year actually finding fair use. Though sufficient for purposes of basic statistical analysis, this is a surprisingly low number of opinions for such an important area of copyright law, particularly one that has received so much academic attention.⁶³ It is all the more surprising in light of Federal Judicial Center data that suggests that a steady average of approximately 2,000 copyright infringement complaints were filed per year in federal district courts during the same period.⁶⁴ A number of

⁶⁰ See, e.g., *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.”).

⁶¹ Cf. RONALD DWORKIN, TAKING RIGHTS SERIOUSLY 15-16 (1978); Roscoe Pound, *Mechanical Jurisprudence*, 8 COLUM. L. REV. 605 (1908).

⁶² The Appendix discusses the “substantial use” standard used to filter the opinions. In short, the data set included all opinions from the period sampled that cited to the Section 107 test and referenced at least two factors from the test.

⁶³ There were typically far more law review articles than actual court opinions on fair use in each of the years sampled. A search of the Westlaw tp-all database for the years 1990-2005, using the search string *ti("fair use") & da([year])*, shows that the ratio of law review articles in that database containing the phrase “fair use” in their titles to actual court opinions on fair use was about 2.4 during that period. For the years 2000-2005, the ratio was 3.3.

⁶⁴ See Inter-university Consortium for Pol. & Soc. Res., Federal Court Cases: Integrated Data Base, 2003, ICPSR Study No. 4026 (2004); Inter-university Consortium for Pol. & Soc. Res., Federal Court Cases: Integrated Data Base, 2002, ICPSR Study No. 4059 (2003); Inter-university Consortium for Pol. & Soc. Res., Federal Court Cases: Integrated Data Base, 2001, ICPSR Study No. 3415 (2002); Inter-university Consortium for Pol. & Soc. Res., Federal Court Cases: Integrated Data Base, 1970-2000, ICPSR Study No.

factors may account for the paucity of reported fair use opinions, the most obvious being that many fair use disputes may never reach the courts.⁶⁵ In any event, it is important to recognize that this study is concerned with a small, but crucially important region in the larger world of fair use: its case law. The data can tell us a great deal about the case law and how the four-factor test operated in it. But because of the fundamental problem of selection bias, we cannot reasonably ask the data to do more.⁶⁶

1. Distribution of the Opinions by Year and Venue

Figure 2 sets out the distribution by year of the district and circuit opinions studied. Though the data show no clear turning point, the average number of district court fair use opinions per year shifted substantially through the course of the late 1980's and early 1990's. This shift may reflect at least in part the appearance in the fair use case law of disputes involving computer software, the first fair use opinion involving which appeared in 1988,⁶⁷ and internet technology, the first fair use opinion involving which appeared in 1993.⁶⁸ The Supreme Court has addressed Section 107 in seven opinions from four cases, all of them falling within the period studied: *Sony v. Universal City Studios*⁶⁹ (1984), *Nation v. Harper & Row*⁷⁰ (1985), *Stewart v. Abend*⁷¹ (1990), and *Campbell v. Acuff-Rose*⁷² (1994). *Sony* and *Harper & Row* each produced one dissent,⁷³ while *Campbell* produced a brief concurrence.⁷⁴

8429 (last updated 2005). According to this database, for the years 1978 up to and including 2003, the mean number of filings per year under the Nature of Suit category “820 Copyright” was 1990.67 (SD=6.83).

⁶⁵ Even while fair use concerns may pervade the public sphere (and legal-academic commentary on it), many fair use controversies may never take the form of a filed complaint, as potential fair uses are “chilled” by the mere threat of litigation. Empirical work supports this explanation. See MARJORIE HEINS & TRICIA BECKLES, WILL FAIR USE SURVIVE? FREE EXPRESSION IN THE AGE OF COPYRIGHT CONTROL (2005). Indeed, the very unpredictability of fair use outcomes may lead defendant's to settle, whether before or after litigation has been joined, particularly when the defendant need only make narrow alterations in its speech in order to avert the risk of broad injunctive relief and the imposition of damages. For a theoretical account of the cost-benefit analysis underlying the pleading of a fair use defense, see generally Thomas Cotter, *Fair Use and Copyright Overenforcement* (working paper on file with author).

⁶⁶ See generally Ahmed E. Taha, *Data and Selection Bias: A Case Study*, 75 U.M.K.C. L. REV. 171 (2006) (discussing the effects of selection bias on the results of empirical studies of whether judges' political orientations affect litigation outcomes). See also David A. Hoffman & Alan J. Izenman, *Docketology, District Courts, and Doctrine* (working paper on file with author).

⁶⁷ See *Telerate Systems, Inc. v. Caro*, 689 F. Supp. 221 (S.D.N.Y. 1988).

⁶⁸ See *Playboy Enters. v. Frena*, 839 F. Supp. 1552 (M.D. Fla. 1993).

⁶⁹ *Sony Corp. of America v. Universal City Studios, Inc.*, 464 U.S. 417 (1984).

⁷⁰ *Harper & Row Publishers, Inc. v. Nation Enterprises*, 471 U.S. 539 (1985).

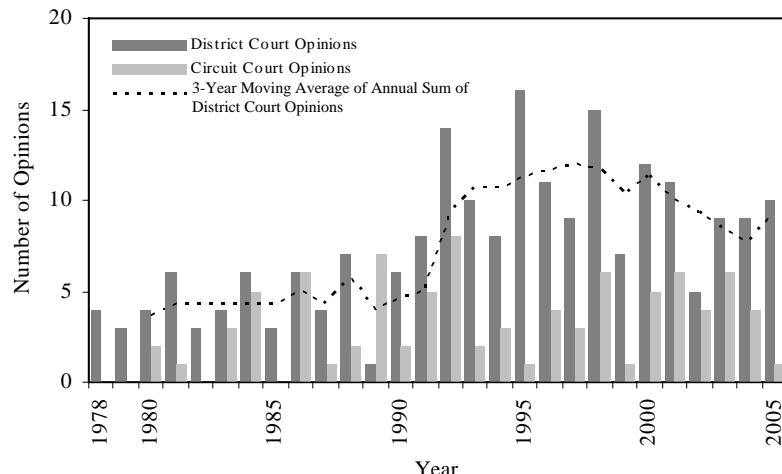
⁷¹ 495 U.S. 207 (1990).

⁷² *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569 (1994)

⁷³ *Harper & Row*, 471 U.S. at 580 (Brennan, J., dissenting); *Sony*, 464 U.S. at 457 (Blackmun, J., dissenting).

⁷⁴ *Campbell*, 510 U.S at 596 (Kennedy, J., concurring).

FIGURE 2
DISTRIBUTION OF DISTRICT AND CIRCUIT COURT OPINIONS BY YEAR



The district and circuit courts of the Second and Ninth Circuits dominated the sample⁷⁵ across the twenty-eight year period⁷⁶ in two closely related ways. First, their courts contributed the great bulk of the opinions sampled. Table 1 reports that Second Circuit courts accounted for 38.6% of the circuit courts opinions and 35.1% of the district court opinions, while Ninth Circuit courts were responsible for 28.4% of the circuit court opinions and 18.0% of the district court opinions. Percentages for all other circuits, even the Seventh, were in the single digits. At the district court level, the Southern District of New York alone accounted for 31.3% of the district court opinions, with the Northern District of California next at 7.6%.

Second, the data show that fair use opinions from courts of the Second and Ninth Circuits exerted a great deal of influence on fair use opinions outside of those circuits, much more than is generally thought. Specifically, data on the import and export of fair use case citations between the various circuits in the opinions studied demonstrate that the Second and the Ninth Circuits both enjoyed substantial trade surpluses. For example, circuit and district court opinions from outside of the Second Circuit cited to an average of 1.55 circuit court cases from the Second Circuit per fair use analysis, and those from

⁷⁵ I use the term “sample” because though this project sought to study the entire population of federal opinions that applied the four factor test from 1978 through 2005, I cannot be sure that I collected all such opinions.

⁷⁶ There was no significant variation over time in the distribution of opinions by venue or in the degree to which opinions from courts of the Second and Ninth Circuit were cited by courts outside of those circuits. It is notable, however, that 35 (or 92%) of the 38 opinions sampled from the district courts of the Ninth Circuit dated from after 1990. This does not appear to reflect the rise of computer software and internet fair use case law, however, as only six of these 35 opinions involved computer software or internet facts.

TABLE 1
DISTRIBUTION OF CIRCUIT AND DISTRICT COURT OPINIONS

Circuit	Circuit Court Opinions		District Court Opinions				
	N	%	N	%	District	N	%
1	1	1.1	8	3.8	D. Mass.	7	3.3
					D.N.H.	1	0.5
2	34	38.6	74	35.1	N.D.N.Y.	3	1.4
					E.D.N.Y.	2	0.9
					S.D.N.Y.	66	31.3
					W.D.N.Y.	3	1.4
3	1	1.1	9	4.3	D.N.J.	2	0.9
					E.D. Pa.	5	2.4
					W.D. Pa.	1	0.5
					D. Del.	1	0.5
4	3	3.4	12	5.7	D. Md.	4	1.9
					E.D. Va.	5	2.4
					M.D.N.C.	1	0.5
					W.D. Va.	2	0.9
5	4	4.6	8	3.8	N.D. Tex.	2	0.9
					E.D. Tex.	1	0.5
					S.D. Tex.	4	1.9
					W.D. Tex.	1	0.5
6	8	9.1	12	5.7	E.D. Ky.	1	0.5
					E. D. Mich.	6	2.8
					W.D. Mich.	2	0.9
					N.D. Ohio	1	0.5
					E. D. Tenn.	1	0.5
					M.D. Tenn.	1	0.5
7	3	3.4	12	5.7	N.D. Ill.	11	5.2
					S.D. Ill.	1	0.5
8	2	2.3	10	4.7	D. Minn.	8	3.8
					W.D. Mo.	1	0.5
					D. Neb.	1	0.5
9	25	28.4	38	18.0	N.D. Cal.	16	7.6
					C.D. Cal.	12	5.7
					S.D. Cal.	5	2.4
					D. Nev.	2	0.9
					D. Ore.	3	1.4
10	---	---	7	3.3	D. Colo.	2	0.9
					D. Kansas	4	1.9
					N.D. Okla.	1	0.5
11	6	6.8	14	6.6	M.D. Fla.	2	0.9
					S.D. Fla.	3	1.4
					N.D. Ga.	9	4.3
DC	1	1.1	7	3.3	D.D.C.	7	3.3
Total	88		211			211	

outside of the Ninth Circuit cited to an average of .68 circuit court cases from the Ninth Circuit per fair use analysis. No other circuits boasted nearly as much influence outside of their borders. The next most influential circuit after the Ninth was the Fifth, with .14 Fifth Circuit cases cited per fair use analysis by circuit and district courts outside of the Fifth Circuit—and this is largely the result of the authority of pre-1981 Fifth Circuit case law for courts of the Eleventh Circuit.⁷⁷ In fact, by this measure, the district courts of the S.D.N.Y. were more influential than any circuit court other than the Second and the Ninth and were nearly as influential as the Ninth Circuit. Circuit courts outside of the Second Circuit cited to an average of .59 S.D.N.Y. cases per fair use analysis, while district courts outside of the Second Circuit cited to an average of .60 S.D.N.Y. cases per fair use analysis. Ultimately, circuit and district courts of the Second, Fifth, and Ninth were the only courts that cited on average more often to opinions from their own circuit court than to opinions from the circuit courts of either the Second or the Ninth Circuits.

Thus, when we speak of modern U.S. fair use case law, we are speaking primarily of the 122 opinions generated by four courts—the Supreme Court, the Second Circuit, the Ninth Circuit, and the Southern District of New York—and the progeny of these opinions in the other federal courts. This makes the detection of intercircuit variation in our fair use case law a hazardous enterprise. Nevertheless, as I will discuss in Part IV, the data suggest that, when taking into consideration a variety of other factors, the Second Circuit stands out as a relatively fair use-friendly circuit.⁷⁸

2. Distribution of District Court Opinions by Posture

Table 2 reports the distribution of postures in the district court opinions. More than half of the opinions addressed a motion or cross-motion for summary judgment, and of these 121 opinions, 86% granted the motion or one of the cross-motions. This supports the conventional wisdom that courts regularly resolve fair use issues at the summary judgment stage.⁷⁹

More interesting is the distribution of the posture of district court opinions across time. Figure 3 shows a moving average across time of the proportion of district court opinions that took the form of preliminary injunction, summary judgment, and bench trial opinions, respectively. In 1994, there was a substantial drop-off in the proportion per year of bench trial opinions, so that up

⁷⁷ See *Bonner v. City of Prichard*, 661 F.2d 1206, 1209 (11th Cir. 1981) (holding that cases decided by the Fifth Circuit Court of Appeals prior to October 1, 1981 are the law of the Eleventh Circuit).

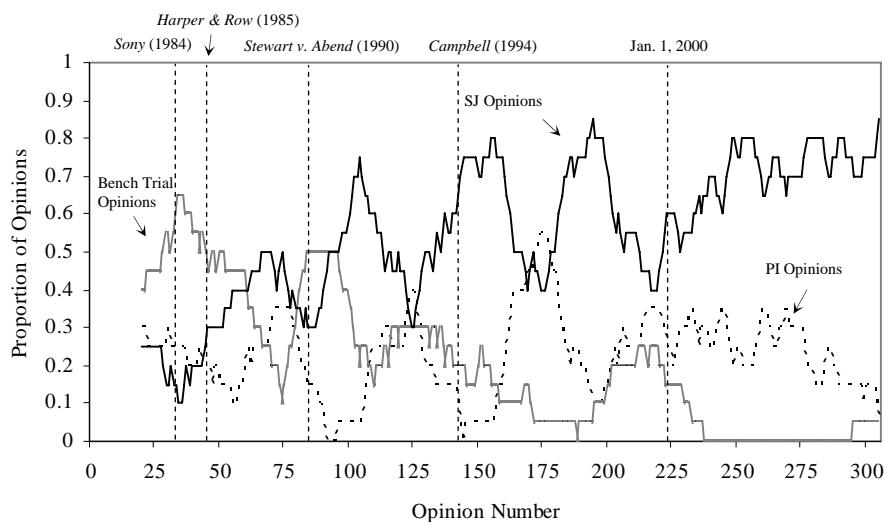
⁷⁸ See *infra* note ___ and accompanying text.

⁷⁹ See, e.g., *Castle Rock Entertainment, Inc., v. Carol Publishing Group, Inc.* 150 F.3d 132, 136 (2d.Cir. 1998) (“Though recognizing that fair use is a ‘mixed question of law and fact,’ courts regularly resolve fair use issues at the summary judgment stage where there are no genuine issues of material fact.”).

TABLE 2
DISTRIBUTION OF DISTRICT COURT OPINIONS BY POSTURE
(“P Found FU” denotes the proportion of opinions which found fair use; “P Found No FU” denotes the proportion of opinions which found no fair use)

Posture	N	%	Found FU	Found No FU
Preliminary Injunction	50	23.7	.300	.700
SJ-Plaintiff	24	11.4	---	.875
SJ-Defendant	38	18.0	.763	---
Cross-SJs	59	28.0	.389	.525
Bench Trial	37	17.5	.324	.676
MTD	2	1.0	.000	1.000
JNOV Motion	1	0.5	.000	1.000
Total	211		.389	.588

FIGURE 3
20-OPIIN MOVING AVERAGE OF THE PROPORTION OF DISTRICT COURT PRELIMINARY INJUNCTION,
BENCH TRIAL, AND SUMMARY JUDGMENT OPINIONS



to and including 1993, the sample included, in absolute numbers, an average of 1.9 bench trial opinions per year, and .5 bench trial opinions per year thereafter. Indeed, for the seven year period from 1999 to 2005, the sample included only one district court bench trial opinion.⁸⁰ At the same time, the data indicate a fairly dramatic increase beginning in the early to mid ‘90s in the proportion of opinions which engaged in summary adjudication of the fair use defense.

Though these results are consistent with recent work on the “vanishing trial” in American courts,⁸¹ they nevertheless remain something of a mystery. The drop-off in bench trials coincides with the Supreme Court’s issuance of its landmark opinion in *Campbell v. Acuff-Rose*⁸² in March, 1994. The district court in *Campbell* had granted summary judgment to the defendant, finding that the musical group 2 Live Crew’s appropriation of certain elements of Roy Orbison’s song “Pretty Woman” in its parody of the song constituted a fair use.⁸³ In reversing and remanding, the Sixth Circuit criticized the district court for giving insufficient weight to the presumption established by the *Sony* Court a decade earlier that “every commercial use . . . is presumptively unfair.”⁸⁴ In itself reversing and remanding, the *Campbell* Court emphatically rejected the *Sony* presumption and explicitly demoted the commerciality of the defendant’s use to merely one issue among others that a court may consider as part of its factor one analysis of the “purpose and character” of the defendant’s use.⁸⁵ The Court then remanded for further fact-finding under factors three and four of Section 107.⁸⁶ The mystery is thus that the reported drop-off in bench trial opinions and increase in summary judgment opinions coincided with a Supreme Court opinion that not only remanded a summary judgment ruling, but abrogated a key presumption facilitating summary adjudication of the fair use defense.

In addition to the more general vanishing trial phenomenon, there may be two copyright-specific explanations for this coincidence—though, admittedly, neither of them is fully satisfactory. First, though the Supreme Court clearly abrogated the *Sony* presumption in its *Campbell* majority opinion, the data show, as we will see below, that lower courts continued to cite and apply that presumption. *Campbell’s* authority was and remains far from absolute. Second, the drop-off in the proportion of bench trial opinions beginning in the mid-1990s may reflect the influence not of the Supreme Court, but of the circuit courts. The early 1990’s saw a wave of circuit court affirmances of summary

⁸⁰ Schiffer Publ’g, Ltd. v. Chronicle Books, No. 03-4962, 2004 U.S. Dist. LEXIS 23052 (E.D. Pa. 2004).

⁸¹ See, for example, the articles collected in 1 J. EMPIRICAL LEGAL STUD. 459-984 (2004).

⁸² 510 U.S. 569 (1994).

⁸³ See *Acuff-Rose Music, Inc. v. Campbell*, 754 F.Supp. 1150 (M.D. Tenn. 1991).

⁸⁴ *Acuff-Rose Music, Inc. v. Campbell*, 972 F.2d 1429, 1436-37 (6th Cir. 1992) (*quoting* Harper & Row Publishers, Inc. v. Nation Enterprises, 471 U.S. 539, 562 (1985) (*quoting* Sony Corp. of America v. Universal City Studios, Inc., 464 U.S. 417, 451 (1984))).

⁸⁵ See *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 583-84 (1994).

⁸⁶ See *id.* at 589.

judgment decisions in fair use cases.⁸⁷ This may have emboldened district court judges and litigants to turn more readily to summary adjudication.

3. Distribution of the Opinions by Subject Matter

Notwithstanding scholarly emphasis on fair use and new media, traditional two-dimensional non-virtual print media have dominated and continue to dominate the facts of American fair use case law. Over half (52.9%) of all of the opinions studied (and 50.7% of the district court opinions) addressed facts in which both parties were engaged in the non-virtual print medium.⁸⁸ More specifically, 36.6% of the opinions (and 31.3% of the district court opinions) addressed facts in which both parties were engaged strictly in the medium of non-virtual text.⁸⁹ Opinions addressing facts involving computer software and/or internet technology began to play a substantial, though far from leading, role in the case law beginning in 1988. For the period 1988-2005, 21.6% of the opinions (and 22.6% of the district court opinions) addressed facts involving computer software and/or the internet. Opinions addressing facts involving video, broadly defined as moving images in television, motion picture, or other form, constituted 20.6% of the opinions, while opinions involving music made up only 6.2% of those studied. Finally, 84.6% of the opinions addressed facts in which both parties' works appeared in the same medium. Where a shift in medium did occur, the most common was from print to video or vice-versa, which was reported in thirteen (or 4.2%) of the opinions.

Unsurprisingly, First Amendment concerns figured prominently in the opinions.⁹⁰ Twenty-five percent of the opinions invoked the First Amendment or free speech concerns more generally through the course of the opinion, though not all did so within the fair use analysis itself. Perhaps also unsurprisingly, the proportion of opinions addressing the First Amendment increased with the authority of the court. Twenty percent of the district court

⁸⁷ See, e.g., National Rifle Ass'n of Am. v. Handgun Control Fed'n, 15 F.3d 559 (6th Cir. 1994); Twin Peaks Prods. v. Publ'n Int'l, Ltd., 996 F.2d 1366 (2d Cir. 1993); Arica Inst., Inc. v. Palmer, 970 F.2d 1067 (2d Cir. 1992); Wright v. Warner Books, Inc., 954 F.2d 731 (2d Cir. 1991); Wright v. Warner Books, Inc., 953 F.2d 731 (2d Cir. 1991); BellSouth Advertising & Publishing Corp. v. Donnelley Info. Publishing, 933 F.2d 952 (11th Cir. 1991); Cable/Home Communication Corp. v. Network Productions, Inc., 902 F.2d 829 (11th Cir. 1990); McGowan v. Cross, No. 92-1480, 1993 U.S. App. LEXIS 9134 (4th Cir. 1993).

⁸⁸ Here, I broadly define the print medium as a medium consisting of two-dimensional textual or graphic works on paper, canvas, or a similar substrate, but not on a computer screen; the definition also excludes computer software code. Sixty-four percent of the opinions, and 64% of the district court opinions specifically, addressed facts in which at least one party was engaged in the print medium.

⁸⁹ Forty-five percent of all opinions, and 43% of district court opinions, addressed facts in which at least one party was engaged in text.

⁹⁰ See generally Neil Netanel, *Locating Copyright Within the First Amendment Skein*, 54 STAN. L. REV. 1 (2001); Rebecca Tushnet, *Copy This Essay: How Fair Use Doctrine Harms Free Speech and How Copying Serves It*, 114 YALE L.J. 535 (2004).

opinions addressed in some way the First Amendment, while 34% of the circuit court opinions and 43% of the Supreme Court opinions did so.⁹¹

B. Reversal, Dissent, and Appeal Rates

The conventional wisdom is that “reversals and divided courts are commonplace”⁹² in the fair use case law. “The field is littered,” we are told, “with the corpses of overturned opinions.”⁹³ This is certainly true of the Supreme Court fair use case law. In *Sony*, *Harper & Row*, and *Campbell*, the Supreme Court reversed the circuit court’s reversal of the district court.⁹⁴ This, together with the fact that the Court was divided five-four in *Sony*⁹⁵ and six-three in *Harper & Row*,⁹⁶ may account for the general belief that our fair use case law is especially unstable.

The data on the case law below the Supreme Court run contrary to the conventional wisdom, however. The sample consisted of 88 circuit court opinions, of which four were concurrences and thirteen were dissents, from 71 cases. Of the 71 majority opinions, 24 reversed the district court’s fair use holding (for a reversal rate of 33.8%) and ten met with dissents on the fair use issue (for a dissent rate of 14.1%).⁹⁷ These results are not substantially different from recent estimates of overall circuit court reversal rates (for example, 32% across all circuits for the period 1980-2002⁹⁸) and dissent rates (for example, 9.4% across all circuits for the period 1970 to 1988⁹⁹)—and the

⁹¹ Here, the literal methods of the “content analysis” approach to the case law do not serve us particularly well, or at least do not produce interesting data. Of the 69 opinions that referenced the First Amendment, 40.6% found in favor of the defendant, as against the 237 opinions that did not reference the First Amendment, of which 42.2% found in favor of the defendant. Nor was it possible reliably to conduct a word-count analysis of courts’ discussion of the First Amendment, as discussions of fair use and the First Amendment tended to be quite discursive in nature. Here, then, the leading cases approach seems to be the only workable method of analysis currently available to us. See, e.g., Netanel, *supra* note 90.

⁹² Pierre N. Leval, *Toward a Fair Use Standard*, 103 HARV. L. REV. 1105, 1106-07 (1990).

⁹³ Weinreb, *supra* note 19, at 1137.

⁹⁴ See *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569 (1994), *rev’g* 972 F.2d 1429 (6th Cir. 1992), *rev’g* 754 F.Supp. 1150 (M.D. Tenn. 1991); *Harper & Row Publishers, Inc. v. Nation Enterprises*, 471 U.S. 539 (1985), *rev’g* 723 F.2d 195 (2d Cir. 1983), *rev’g* 501 F.Supp. 848 (S.D.N.Y. 1980); *Sony Corp. of America v. Universal City Studios, Inc.*, 464 U.S. 417 (1984), *rev’g* 659 F.2d 963 (9th Cir. 1981), *rev’g* 480 F.Supp. 429 (C.D. Cal. 1979).

⁹⁵ See *Sony*, 464 U.S. at 419, 457.

⁹⁶ See *Harper & Row*, 471 U.S. at 541, 579

⁹⁷ Two cases produced multiple dissents from en banc hearings. See *Princeton Univ. Press v. Michigan Document Servs.*, 99 F.3d 1381, at 1393 (6th Cir. 1996) (Boyce, C.J., dissenting); *id.* at 1394 (Merritt, J., dissenting); *id.* at 1397 (Ryan, J., dissenting); *New Era Publications International v. Henry Holt, Co.*, 884 F.2d 659, 662 (2d Cir. 1989) (Newman, J., dissenting)..

⁹⁸ See Kevin M. Scott, *Understanding Judicial Hierarchy: Reversals and the Behavior of Intermediate Appellate Judges*, 40 LAW & SOC’Y REV. 163, 176 (2006). See also Burton Atkins, *Interventions and Power in Judicial Hierarchies: Appellate Courts in England and the United States*, 24 LAW & SOC’Y REV. 71 (1990) (comparing reversal rates in U.S. appellate courts to those in English appellate courts).

⁹⁹ SEE DONALD R. SONGER ET AL., CONTINUITY AND CHANGE ON THE UNITED STATES COURTS OF APPEALS 105 (2000) (Table 5.1) (reporting a 9.43% dissent rate from a sample of all published circuit court cases from 1970 to 1988). See also Sean Farhang & Gregory Wawro, *Institutional Dynamics on the U.S. Court of*

circuit court reversal rate in fair use cases is well below estimates of Federal Circuit reversal rates in patent cases (for example, 47.3% for the period January, 1998 through April, 2000¹⁰⁰). Nor is the appellate data particularly exceptional when viewed from the perspective of district court opinions. Of the 211 district court opinions, 53 were appealed (for an unexceptional appeal rate of 25.1%¹⁰¹), with 35 of the 211 (or 16.6%) affirmed and 18 (or 8.5%) reversed.¹⁰² Overall, these data support the unexpected, if also perhaps somewhat uninspiring, finding that our fair use case law, at least outside of the cases that reached the Supreme Court (and our casebooks), has not been marked by especially high reversal, dissent, or appeal rates.

C. Fair Use Win Rates

How often have courts actually found fair use in our Section 107 case law? The fair use win rate results are among the most curious in this study. I first set out the basic results and then attempt some explanations.

Tables 3 and 4 report the fair use win rate results in the district court and circuit court opinions, respectively. More specifically, Table 3 reports the proportion of unreversed district court opinions, grouped by circuit and posture, that either found fair use or, in the case of summary judgment motions by the plaintiff, found no fair use. Overall, 30.4% of the preliminary injunction opinions found fair use, while 24.1% of the bench trial opinions did so. Among summary judgment opinions 86.4% of the opinions that addressed a plaintiff's uncrossed motion for summary judgment granted that motion, while 75.7% of the opinions that addressed a defendant's uncrossed motion for summary judgment granted that motion. The parties' win rates were sharply lower for cross-motions for summary judgment. As Table 4 reports, the fair use win rate results in the circuit court majority opinions are roughly comparable to those in the district court opinions, but as a matter of statistical significance, none of the circuit court results were significantly different from 50 percent.

Considering first the district court results, it should not be surprising that a high proportion of district court opinions addressing an uncrossed summary

Appeals: Minority Representation Under Panel Decision Making, 20 J.L. ECON. & ORG. 299, 306 (2004) (“Empirically, federal appellate panels are overwhelmingly unanimous, with dissent rates aggregated across all circuits averaging approximately 6% to 8%, varying somewhat with respect to issue area.”).

¹⁰⁰ Christian A. Chu, *Empirical Analysis of the Federal Circuit’s Claim Construction Trends*, 16 BERKELEY TECH. L.J. 1075, 1098 (2001). See also Stephen P. Swinton & Adam A. Welland, *Patent Injunction Reform and the Overlooked Problem of ‘False Positives’*, 70 BNA PAT. TRADEMARK & COPYRIGHT J. 337, 338 (2005) (“Depending on the analysis employed and period examined, reversal rates for trial court judgments in patent infringement cases range from 30 to more than 60 percent.”).

¹⁰¹ See Theodore Eisenberg, *Appeal Rates and Outcomes in Tried and Nontried Cases: Further Exploration of Anti-Plaintiff Appellate Outcomes*, 1 J. EMPIRICAL LEGAL STUD. 659 (2004).

¹⁰² Of the 18 district court opinions that were reversed, nine found in favor of the plaintiff and nine found in favor of the defendant.

TABLE 3
DISPOSITION OF UNREVERSED DISTRICT COURT FAIR USE OPINIONS BY CIRCUIT AND POSTURE
("FU" denotes the proportion of opinions which found fair use; "No FU" denotes the proportion of opinions which found no fair use)

Circuit	Distribution by		Preliminary		Bench Trial		SI Plaintiff		SI Defendant		SI Cross		
	Circuit	N	% Circuit	N	FU	N	FU	N	No FU	N	FU	N	No FU
1	8	4.2		2	.500	3	.667	---	---	2	1.000	1	.000
2	64	33.7		12	.500	9	.000	6	.667	18	.778	19	.421
3	9	4.7		4	.000	2	.000	---	---	2	1.000	1	1.000
4	12	6.3		1	1.000	1	1.000	1	1.000	4	.750	5	.800
5	7	3.7		1	1.000	5	.200	---	---	---	---	1	.000
6	11	5.8		1	.000	1	.000	3	.667	1	1.000	5	.600
7	10	5.3		4	.000	1	.000	1	1.000	---	---	4	.750
8	9	4.7		2	.000	1	.000	1	1.000	2	.500	3	1.000
9	34	17.9		11	.181	2	1.000	4	1.000	6	.500	11	.273
10	7	3.7		3	.667	1	.000	1	1.000	1	1.000	1	.000
11	12	6.3		3	.000	2	.500	4	1.000	---	---	3	1.000
D.C.	7	3.7		2	.500	1	.000	1	1.000	1	1.000	2	.000
Total	190*			46		29		22		37		56	
Circuit-Wide:				304		241		.864		.757		.536	
													.375

*Three unreversed district court opinions are not represented in this table. None of them found fair use. See Chicago Sch. Reform Bd. of Trustees v. Substance Inc., 79 F. Supp. 2d 919 (N.D. Ill. 2000) (motion to dismiss); Int-Elect Eng'g, Inc. v. Clinton Harley Corp., 1993 U.S. Dist. LEXIS 11510 (S.D. Cal. 1993) (same); Roy Export Co. v. Columbia Broadcasting System, Inc., 503 F. Supp. 1137 (S.D.N.Y. 1980) (motion for I&NOV).

[Vol. 001: 1]

TABLE 4
DISTRIBUTION AND DISPOSITION OF UNREVERSED CIRCUIT COURT MAJORITY OPINIONS
(“FU” denotes the proportion of opinions which found fair use; “No FU” denotes the proportion of opinions which found no fair use)

Circuit	Distribution by Circuit		Preliminary Injunction		Bench Trial		SJ Plaintiff		SJ Defendant		SJ Cross	
	N	%	N	FU	N	FU	N	No FU	N	FU	N	No FU
1	1	1.6	---	---	---	---	1	1.000	---	---	---	---
2	24	37.5	5	.400	7	.143	2	.500	4	.750	6	.167
3	1	1.6	---	---	---	---	---	---	---	1	1.000	.000
4	3	4.7	---	---	1	1.000	---	---	1	1.000	1	.000
5	2	3.1	---	---	1	1.000	---	---	1	1.000	0	000
6	3	4.7	---	---	1	1.000	---	---	2	.500	500	
7	2	3.1	---	---	---	---	1	.000	1	.000	---	---
8	2	1.6	---	---	---	---	---	---	1	.000	.000	
9	21	32.8	8	.375	2	.500	1	1.000	3	.333	7	.286
10	---	---	---	---	---	---	---	---	---	---	---	---
11	5	7.8	1	1.000	1	.000	2	1.000	---	1	1.000	.000
D.C.	1	1.6	1	.000	---	---	---	---	---	---	---	---
Total	64	15	13	.400	385	.667	10	.600	20	.350	.550	
Circuit-Wide:												

*Three unreversed circuit court opinions are not represented in this table. See *Mulcahy v. Cheetah Learning LLC*, 386 F.3d 849 (8th Cir. 2004) (finding certain uses at issue fair and others unfair); *TY, Inc. v. Publits Int'l*, 292 F.3d 512 (7th Cir. 2002) (same); *Ringgold v. Black Entertainment TV*, 126 F.3d 70 (2d Cir. 1997) (remanding on the fair use issue for further finding of facts)

judgment motion granted the motion. As an empirical matter, these results are consistent with similar results from a study of the multifactor test for consumer confusion in U.S trademark cases.¹⁰³ As a theoretical matter, parties are less likely to incur the costs of filing a motion for summary judgment where the motion lacks merit, opposing parties are more likely to file a cross-motion where they have a reasonably strong case, and district court judges are more likely to write published opinions when they grant rather than deny motions for summary judgment.

This still leaves open the question of the exceptionally low fair use win rates in district court preliminary injunction and bench trial opinions. The Priest-Klein “selection hypothesis” predicts that, given various conditions, plaintiff win rates at trial should approach 50 percent, largely because it is only the close cases that survive settlement—or summary adjudication.¹⁰⁴ Evidence of the 50 percent hypothesis has been mixed at best.¹⁰⁵ The fundamental assumption limiting the 50 percent hypothesis, and the data that has been mustered to support it, is that the parties must have equal stakes in the litigation for it to be true.¹⁰⁶ Yet, as William Landes has argued, intellectual property plaintiffs tend to be parties with higher stakes because they face the risk that an adverse judgment will limit or extinguish their rights.¹⁰⁷ We would therefore expect intellectual property plaintiffs to settle even the near-close cases, Landes argues, with the result that plaintiff win rates in intellectual property cases that make it to trial should exceed 50 percent by a comfortable margin. Landes finds strong support for this hypothesis in Federal Judicial Center data showing, for example, a 73% copyright plaintiff trial win rate in federal district courts for the years 1978-2000 as against a 48% plaintiff trial win rate for civil trials generally.¹⁰⁸

The high plaintiff win rates in the district court preliminary injunction and bench trial opinions are certainly consistent with Landes’ more general hypothesis and the evidence he presents in support of it. Yet the Landes hypothesis runs counter to the conventional view among commentators (including myself) of plaintiffs in fair use litigation, if not of copyright owners

¹⁰³ See Barton Beebe, *An Empirical Study of the Multifactor Tests for Trademark Infringement*, 95 CAL. L. REV. 1596-98 (2006).

¹⁰⁴ See George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1 (1984).

¹⁰⁵ See generally Theodore Eisenberg, *Testing the Selection Effect: A New Theoretical Framework with Empirical Tests*, 19 J. LEGAL STUD. 337 (1990); Daniel Kessler, Thomas Meites & Geoffrey Miller, *Explaining Deviations from the Fifty-Percent Rule: A Multimodal Approach to the Selection of Cases for Litigation*, 25 J. LEGAL STUD. 233 (1996).

¹⁰⁶ See Priest & Klein, *supra* note 104, at 24-29.

¹⁰⁷ See William M. Landes, *An Empirical Analysis of Intellectual Property Litigation: Some Preliminary Results*, 41 HOUS. L. REV. 749, 772 (2004).

¹⁰⁸ *Id.* at 774. The Federal Judicial Center data is highly suspect. See generally Theodore Eisenberg & Margo Schlanger, *The Reliability of the Administrative Office of the U.S. Courts Database: An Initial Empirical Analysis*, 78 NOTRE DAME L. REV. 1455 (2003). See also Beebe, *supra* note 103, at 1652-64.

more generally.¹⁰⁹ This view holds that in our “clearance culture,”¹¹⁰ owners of valuable copyrightable expression quite rationally tend to be very aggressive litigants, and not simply because they expose themselves to the defenses of acquiescence or laches if they are not. More importantly, they have a stake in establishing a *reputation* for being aggressive litigants in order to benefit from the “chilling effects” that that reputation may generate.¹¹¹ These forces may play out largely at the cease and desist stage of fair use disputes,¹¹² but we expect them to play a role in plaintiff’s decision to litigate—and not to settle—as well.

It is important to emphasize that the data set collected for this study cannot tell us whether the plaintiffs in the opinions sampled conform to the conventional view just outlined. But I want to take advantage of this view to open the door to a different and tentative explanation for the low fair use win rates in the district court preliminary injunction and bench trial opinions. This explanation assumes that some defendants who are otherwise committed to defending against a copyright infringement claim on grounds such as copyrightability or substantial similarity may find it relatively inexpensive also to plead a fair use defense, even when the defense may be frivolous or at least very weak in light of the facts. Because conscientious judges will dutifully consider each of the four factors, as Section 107 instructs them to do, even when the outcome of the fair use test is obvious, opinions addressing even essentially extraneous fair use defenses will have come within those sampled for this study. This would drive down overall fair use win rates.

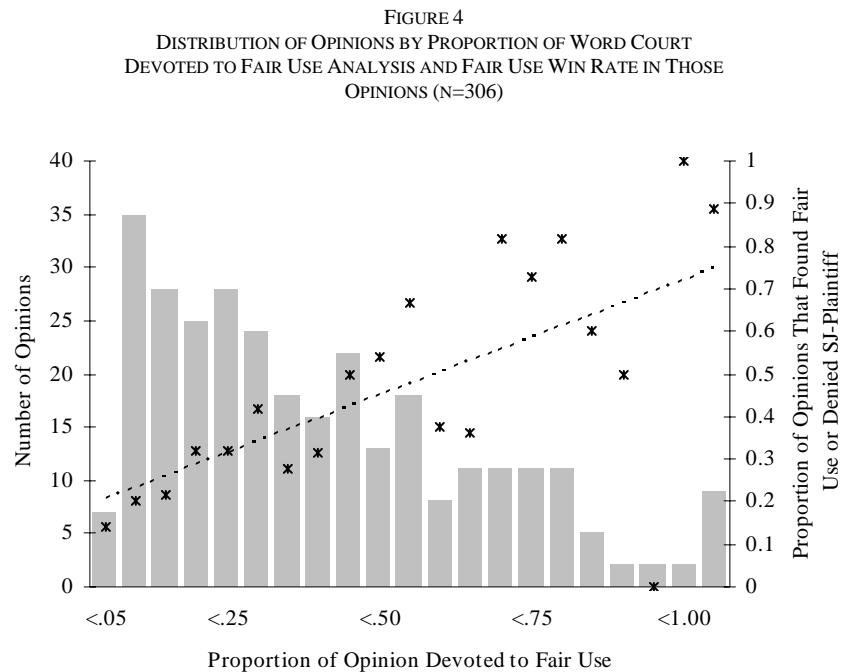
The problem is how objectively to determine whether an opinion addressed a frivolous fair use defense. There is no good way to do so. However, one reasonably workable quantitative index of the weakness or marginality of the defendant’s fair use defense may be the proportion of the opinion—in words—that the judge devoted to the defense. We would expect that, as a general matter, opinions that addressed an extraneous fair use defense would devote a smaller proportion of the opinion to that defense, while opinions that addressed an at least reasonable claim of fair use would spend more time analyzing the defense. Consider, then, Figure 4. This figure presents a histogram of all 306 opinions distributed according to the proportion of the word-count of the opinion devoted to the fair use defense. The figure also reports the defendant

¹⁰⁹ See generally Jane C. Ginsburg, *How Copyright Got a Bad Name for Itself*, 26 COLUM. J.L. & ARTS 61 (2002) (discussing overreaching by copyright owners but defending certain new media copyright reforms). Cf. Jane C. Ginsburg, *Authors and Users in Copyright*, J. COPYRIGHT SOC’Y U.S.A. 1, 1 (1997) (“It has become fashionable, among some thinkers and activists in copyright and related fields, to disparage or to deplore copyright protection.”).

¹¹⁰ See generally PATRICIA AUFDERHEIDE & PETER JASZI, UNTOLD STORIES: CREATIVE CONSEQUENCES OF THE RIGHTS CLEARANCE CULTURE FOR DOCUMENTARY FILMMAKERS (2004), http://www.centerforsocialmedia.org/rock/backgrounddocs/printable_rightsreport.pdf; LAWRENCE LESSIG, FREE CULTURE (2005), <http://www.free-culture.cc/freeculture.pdf>.

¹¹¹ See generally Chilling Effects Clearinghouse, <http://www.chillingeffects.org/>.

¹¹² See *id.*



win rates for opinions falling within each bin of the histogram. As expected, there is a clear and fairly steady positive relation between the proportion of an opinion devoted to the fair use defense and the likelihood that the opinion ultimately ruled in favor of the defendant on the fair use issue. Indeed, if we exclude out the 42 opinions that devoted less than 10% of the opinion to the fair use issue (which are represented in the two leftmost bins), we quickly come to a very respectable defendant win rate in the remaining opinions of 45.5%. To be sure, this is not a comprehensive explanation of the low preliminary injunction and bench trial win rates reported in Table 3, but it may account for at least some of the sub-50 percent nature of the results.

Two final issues deserve mention with respect to the fair use win rate results. First, the far more moderate win rates at the circuit court level, none of them statistically significantly different from 50%, are fully consistent with the underlying logic of the Priest-Klein selection hypothesis. Second, though the data set is not large enough to make strong claims about intercircuit variation in fair use win rates, it is remarkable to note that the district courts of the Seventh Circuit did not find fair use in any of their ten unreversed opinions, while the district courts of the Second Circuit did not find fair use in any of their nine unreversed bench trial opinions. Meanwhile, only two of the eleven unreversed preliminary injunction opinions from the district courts of the Ninth Circuit found fair use.

III. INTERFACTOR ANALYSIS

I turn now to the main focus of this study: how judges use the four-factor test set forth in Section 107 to adjudicate the fair use defense. To begin this analysis, I first consider here how the factor outcomes interacted with each other and with the overall test outcome in the opinions. I also consider the extent to which the factors stampeded to conform to the overall test outcome. This will prepare the ground for the analysis in the next Part of what *subfactor* considerations influenced specific factor outcomes as well as the overall test outcome.

A. Correlation and Regression Analysis

In practice, which factor or factors drive the outcome of the Section 107 test? This question has produced an enormous amount of speculation. Each factor, it seems, has its champions and its detractors,¹¹³ though most courts and commentators assume that, in practice, the outcome of the Section 107 test relies primarily on the outcome of the fourth factor,¹¹⁴ which calls for an economic analysis of the effect of the defendant's use on the market for the plaintiff's work. Indeed, the *Harper & Row* Court called the fourth factor “undoubtedly the single most important element of fair use.”¹¹⁵ The *Campbell* Court subsequently tried to overwrite that dictum,¹¹⁶ but with only limited success, as we will see in the next Part. Commentators tend further to assume that, in practice, the outcomes of the first factor (concerning the purpose and character of the defendant's use) and the fourth factor often coincide, so often, in fact, that several commentators have expressed concern about courts'

¹¹³ On the first factor, see, for example, *On Davis v. Gap, Inc.*, 246 F.3d 152, 171 (2d Cir. 2001) (Leval, J.) (referring to the first factor as “the heart of the fair use inquiry”); *Kelly v. Arriba Soft Corp.*, 77 F.Supp.2d 1116, 1121 (C.D. Cal. 1999) (“The first factor of the fair use test is the most important in this case.”); *Hofheinz v. Discovery Communications, Inc.*, No. 00-3802, 2001 WL 1111970, at *5 (S.D.N.Y. 2001) (“[T]his first and foremost factor strongly favors the defendant.”); Leval, *supra* note 92, at 1116 (“Factor One is the soul of fair use.”). On the second factor, see, for example, *Harper & Row Publishers, Inc. v. Nation Enterprises*, 471 U.S. 539, 553 (1985) (referring to the second factor as “highly relevant to whether a given use is fair”); *Cable/Home Communication Corp. v. Network Cable/Home Communication Corp. v. Network Productions, Inc.*, 902 F.2d 829, 844 (11th Cir. 1990) (“This factor is ‘highly relevant to whether a given use is fair.’ Harper & Row, 471 U.S. at 552-53.”). *But see Dow Jones & Co. v. Board of Trade*, 546 F. Supp. 113, 120 (S.D.N.Y. 1982) (“The nature of the copyrighted work seems to be the least important and most unclear of the four factors enumerated in § 107.”); *Carroll*, *supra* note 11, at ___ (factor two “tends to do little work in swaying the outcome” of the test); *Sag*, *supra* note 11, at 390 (“The nature of the copyrighted work, while fairly objective, nonetheless remains unhelpful in assessing whether an activity is protected by fair use or not because it is overwhelmed by the other factors.”). On the third factor, see, for example, *Compaq Computer Corp. v. Procom Technology, Inc.*, 908 F. Supp. 1409, 1421 (S.D.Tex. 1995) (“The third factor . . . is generally considered the least important factor of the fair use analysis.”).

¹¹⁴ See, e.g., *Nimmer*, *supra* note 14, at 267 n. 25 (“My own opinion is that the fourth factor is the most important.”).

¹¹⁵ *Harper & Row*, 471 U.S. at 566.

¹¹⁶ See *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 578 (1994) (“[A]ll [factors] are to be explored, and the results weighed together, in light of the purposes of copyright.”).

“double counting” of the same considerations under the two factors.¹¹⁷ This concern is based on the nature of the doctrine informing the first and fourth factor analyses. If a court finds that the defendant’s use is “transformative” or “non-commercial” under factor one, and that factor one therefore favors the defendant, a court will also likely find that the defendant’s use, precisely because it is transformative or non-commercial, will not adversely affect the market for the plaintiff’s work. The expected result is that factor four will also favor the defendant.¹¹⁸ The inverse is thought to hold as well, particularly if a court finds that the defendant’s use is “commercial” under factor one. This is because the *Sony* Court established that a use found to be commercial under factor one is presumptively harmful to the market for plaintiff’s work under factor four.¹¹⁹ The *Campbell* Court also tried to overwrite—or at least, modify—this dictum,¹²⁰ again with little success, as we will see in the next Part. As for the middle factors two and three (concerning the nature of the plaintiff’s work and the quantitative and qualitative extent of the defendant’s taking, respectively), commentators tend to regard these, if they regard them at all, as peripheral to the outcome of the test.¹²¹

Correlation analysis provides one simple means to test against the data the conventional wisdom on the relative importance of and interactions among the fair use factors. Table 5 reports the correlation coefficients for the relations between each of the factor outcomes and the overall test outcome as well as among each of the factor outcomes in the 297 dispositive opinions studied, regardless of posture.¹²² These results show that the outcomes of factors one and four very strongly correlated with the test outcome and fairly strongly correlated with each other, while the outcome of factor two correlated weakly, if at all, with the outcome of the test and with the outcomes of the other factors.

¹¹⁷ See Fischer, *supra* note 9, at 1672-1673. See also Jeremy Kudon, *Form Over Function: Expanding the Transformativeness Use Test for Fair Use*, 80 B.U. L. REV. 579, 605 (2000).

¹¹⁸ See, e.g., NXIVM Corp. v. Ross Institute, 364 F.3d 471, 485 (2d Cir. 2004) (“As *Campbell* pointed out, these inquiries, specified in the first and fourth listed factors of § 107, are correlated: the greater the transformative purpose of the secondary use, the less potential purchasers will see it as an alternative means of acquiring the original.”).

¹¹⁹ *Sony Corp. of America v. Universal City Studios, Inc.*, 464 U.S. 417, 451 (1984) (“If the intended use is for commercial gain, [the] likelihood [of significant market harm may be] presumed.”).

¹²⁰ *Campbell*, 510 U.S. at 591.

¹²¹ See *supra* note 113. But see *Leval*, *supra* note 92, at 1122-23 (“This [third] factor has further significance in its bearing on two other factors. It plays a role in consideration of justification under the first factor (the purpose and character of the secondary use); and it can assist in the assessment of the likely impact on the market for the copyrighted work under the fourth factor (the effect on the market.”).

¹²² Note that the sum of the absolute values of the two correlation coefficients shown for each factor does not equal zero because the court could also have found the factor to be neutral, not relevant, a fact issue, or the court’s finding was unclear. For this correlation analysis, each factor outcome is represented with two binary variables: favors a finding of fair use (1=yes, 0=no) and disfavors a finding of fair us (1=yes, 0=no). Thus, if the first variable is coded as one, then the second variable will be coded as zero, and vice-versa. But if the court found the factor to be neutral, irrelevant, or not argued, then both variables were coded as zero.

TABLE 5
CORRELATIONS BETWEEN A FINDING OF FAIR USE AND THE FACTOR OUTCOMES AND AMONG
THE FACTOR OUTCOMES IN 297 DISPOSITIVE OPINIONS

		FU Found / SJ-P Denied	Factor One		Factor Two		Factor Three		Factor Four	
			Favors	Disfavors	Favors	Disfavors	Favors	Disfavors	Favors	Disfavors
	FU Found / SJ-P Denied	1.000								
Factor One	Favors	.747*	1.000							
	Disfavors	-.780*	-.735*	1.000						
Factor Two	Favors	.321*	.310*	-.192*	1.000					
	Disfavors	-.281*	-.087	.267*	-.491*	1.000				
Factor Three	Favors	.646*	.559*	-.496*	.339*	-.190*	1.000			
	Disfavors	-.680*	-.471*	.631*	-.146*	.370*	-.630*	1.000		
Factor Four	Favors	.823*	.713*	-.616*	.364*	-.174*	.637*	-.504*	1.000	
	Disfavors	-.806*	-.557*	.673*	-.205*	.331*	-.474*	.688*	-.722*	1.000

"FU Found / SJ-P Denied" denotes that the court found fair use or otherwise denied the plaintiff's motion for summary judgment on the issue. "Favors" denotes outcomes in which the factor was found to favor a finding of fair use. "Disfavors" denotes outcomes in which the factor was found to disfavor a finding of fair use. * denotes that the coefficient is statistically significant at the .05 level.

TABLE 6
LOGISTIC REGRESSION OF THE OUTCOME OF THE FAIR USE TEST AS A
FUNCTION THE OUTCOMES OF THE SECTION 107 FACTORS IN 297 DISPOSITIVE
OPINIONS

Dependent variable: Fair Use Found (1) / Fair Use Not Found (0)						N=297
						Pseudo R ² =.872
Log likelihood = -25.950						Correctly Classified: 95.6%
	Odds Ratio	Standard Coefficient	Standard Error	P> Z	95% CI	
Factor One	15.005*	2.708*	.683	.000	1.370	4.047
Factor Two	2.702	.994	.772	.198	-.519	2.507
Factor Three	7.339*	1.993*	.743	.007	.538	3.449
Factor Four	39.167*	3.668*	.829	.000	2.043	5.293
constant		.198	.389	.610	-.565	.962

* denotes statistical significance at the .05 level. The "Odds Ratio" suggests, for example, that as factor one shifts from "disfavors fair use" to "other," the odds of the court finding fair use increase by a factor of 15, and as factor one shifts from "other" to "favors fair use," the odds again increase by a factor of 15.

Factor three did better than factor two with respect to the overall test outcome, but the strength of its correlations with the outcomes of factors one and four was not impressive.¹²³

Putting these coefficients in perhaps more easily understood—and starker—terms, the outcome of factor four coincided with the outcome of the overall test in 83.8% of the 297 dispositive opinions while the outcome of factor one coincided with the outcome of the overall test in 81.5% of these same opinions. By comparison, the outcome of factor two coincided with the outcome of the overall test in 50.2% of these opinions. As for the combined influence of factors one and four, in 214 (or 72.1%) of the opinions, factors one and four either both favored or both disfavored fair use. In all but one of these opinions,¹²⁴ the outcome of the fair use test followed the outcome of these two factors. What happened when, if ever, factor one favored (or disfavored) fair use while factor four disfavored (or favored) fair use? Did one of these leading factors consistently trump the other? Factors one and four pointed in opposite directions in only 20 of the opinions. In fourteen of these opinions, the outcome of the test followed the outcome of factor four, while in six, the outcome of the test followed the outcome of factor one. Though hardly conclusive, this breakdown is consistent with the conventional view that factor four exerts the stronger influence on the outcome of the test.

Table 6 reports the results of logistic regression analysis of an overall finding of fair use as a function of the outcomes of each of the four factors in the 297 dispositive opinions. The regression model used here is highly stylized in that, first, it specifies the factor outcomes in the form of trinary explanatory variables coded as favors fair use (1), disfavors fair use (-1), and other (0), and second, it does not include interaction variables. Nevertheless, this very rudimentary model, which I will call the “Simple Model” of the Section 107 test (because I will present a more nuanced model in the next Part), correctly classified 95.6% of the test outcomes. The regression results are consistent with conventional wisdom and with the results of correlation analysis. Controlling for the effects of the other three factors, the first and fourth factors are shown each to exert an enormous amount of influence on the outcome of the test, with the fourth very much in the driver’s seat, while factor two is shown to exert no significant effect on the test outcome.

But perhaps the main teaching of the Simple Model is that we ultimately learn very little from evaluating the Section 107 test at so abstract a level as that of the factor outcomes. It is certainly interesting to observe, now based on empirical evidence, that the outcome of the fourth factor appears to drive the outcome of the test, and that the outcome of the first factor appears also to be

¹²³ There was no significant variation over time in the strength of the correlations.

¹²⁴ See Religious Tech. Ctr. v. Netcom On-Line Commun. Servs., 923 F. Supp. 1231 (N.D. Cal. 1995) (finding no fair use where factors two and three disfavored fair use while factors one and four favored fair use).

highly influential. But of course these findings beg the question of what subfactor considerations lead judges to conclude that the fourth or first factors—or the second or third, for that matter—favored or disfavored fair use. This is the subject of Part IV. Before turning to these intrafactor concerns, however, and as further background for their consideration, I consider two final interfactor questions: to which factors do judges devote the bulk of their written analysis and to what extent do the factors stampede?

B. Word Count Analysis

Because judges tended to conduct their Section 107 analysis in a highly methodical and explicit manner, it was a relatively simple task to establish in the opinions where their discussion of one factor ended and that of another began. This facilitated in turn the analysis of the word count of each opinion devoted to the discussion of a particular factor as a proportion of the word count of the opinion devoted to the overall discussion of the fair use issue. The results of this analysis are reported in Table 7 and Figure 5.

By their terms, none of the factors necessarily calls for more written analysis than the others, and as the standard deviations reported in Table 7 suggest, the opinions varied widely in how much attention each devoted to certain factors. Yet as Table 7 shows, for the 306 opinions, judges tended overall to devote a far greater share of their discussion of the fair use issue to analyses under factors one and four than under factors two and three. If we accept that, in explaining (or defending) their analysis of a legal issue, judges are generally more likely to dedicate a greater share of their explanation to considerations that they deem to be more important to that analysis, then these results lend further support to the finding that, in practice, the fourth and first factors—or more specifically, the subfactors considered under them—drive the outcome of the test.

Figure 5 depicts judges' proportional attention to each factor across time. Most interesting is the story of the first factor. The *Sony* presumption that commercial uses are presumptively unfair likely accounts for the early rise in proportional attention to factor one. This attention then dropped off and factor four rose to prominence until Judge Leval's 1990 Harvard Law Review article *Towards a Fair Use Standard*¹²⁵ urged courts to attend more closely to the question of transformativeness. In 1994, the *Campbell* Court cited to and amplified Judge Leval's teaching,¹²⁶ which very likely gave rise to factor one's golden age in the late 1990s. We then see a fairly precipitous drop-off in attention to factor one, which continued through 2005. What is remarkable is

¹²⁵ See Leval, *supra* note 92, at 1111.

¹²⁶ See *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 578-79 (1994).

2007]

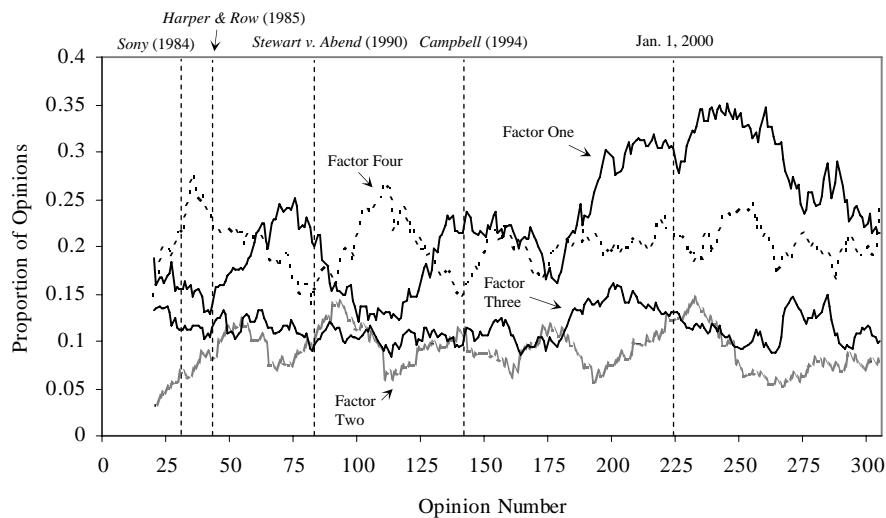
U.S. Copyright Fair Use Opinions

33

TABLE 7
 MEAN PROPORTION OF FAIR USE DISCUSSION DEVOTED
 TO EACH FACTOR IN 306 FAIR USE OPINIONS

	Mean	Std. Dev.	Min/Max
Factor One	.23	.16	.00, .72
Factor Two	.09	.09	.00, .61
Factor Three	.11	.09	.00, .67
Factor Four	.20	.15	.00, .67

FIGURE 5
 20-OPINION MOVING AVERAGE OF THE PROPORTION OF FAIR USE DISCUSSION
 DEVOTED TO EACH FACTOR



not so much that judges' proportional attention to factor one increased after *Sony* and then again after *Campbell*. Rather, it is that this proportional attention repeatedly subsided. First the commerciality inquiry and then the transformativeness inquiry came into and then fell out of fashion. Of this I will have more to say in the next Part.

C. Stampeding

The interfactor correlation coefficients in Table 5 show the degree to which an individual factor outcome correlated with another individual factor outcome. But to what extent did multiple factor outcomes correlate in the opinions? In other words, to what extent did all or most of the factor outcomes tend to stampede in one or the other direction? In a previous study of judges' use of the multifactor test for trademark infringement, I found strong evidence that judges tended to stampede the factors of that test, and argued that this revealed the degree to which judges engaged in "coherence-based reasoning" in applying that test.¹²⁷ The conventional wisdom in copyright law is that judges also stampede the factors of the fair use test. None other than David Nimmer, the author of the authoritative copyright treatise, has argued this in strong terms: "Courts tend first to make a judgment that the ultimate disposition is fair use or unfair use, and then align the four factors to fit that result as best they can. At base, therefore, the four factors fail to drive the analysis, but rather serve as convenient pegs on which to hang antecedent conclusions."¹²⁸ He further asserts, as an empirical matter, that "judges who uphold fair use almost always find that three, if not four, of the factors incline in its favor; judges who deny the fair use defense almost always find that three, if not four, of the factors incline against it."¹²⁹ Other commentators have made similar claims.¹³⁰

Just as it is with respect to reversal rates in the fair use case law, the conventional wisdom with respect to the degree to which courts stampede the fair use factors appears to be based on the two Supreme Court cases *Sony* and *Harper & Row*. And certainly, these cases show stampeding. In *Sony*, the district court found that three (or perhaps four) factors favored fair use,¹³¹ while the Ninth Circuit found that all four factors disfavored fair use.¹³² At the Supreme Court, the five justice majority then found that all four factors favored

¹²⁷ See Beebe, *supra* note 103, at 1615-17.

¹²⁸ Nimmer, *supra* note 14, at 281.

¹²⁹ *Id.* at 280.

¹³⁰ Cf. Sag, *supra* note 11, at 386 ("The current practice of most courts, treating the factors as outcome-determinative as opposed to question-framing, masks a priori assumptions and distorts judicial reasoning."). Indeed, in an earlier working paper version of this project, I read the data to support these claims.

¹³¹ See *Universal City Studios, Inc. v. Sony Corp. of America*, 480 F.Supp. 429 (C.D. Cal. 1979).

¹³² See *Universal City Studios, Inc. v. Sony Corp. of America*, 659 F.2d 963 (9th Cir. 1981).

fair use,¹³³ while the four dissenters found that all four factors disfavored fair use.¹³⁴ *Harper & Row* stampeded back and forth in essentially the same way.¹³⁵

But what of the rest of the fair use case law? Figure 6 reports the distribution of “stampede scores” by posture and outcome among the district court opinions. An opinion’s stampede score is the sum of the factors that the opinion found to favor a finding of fair use minus the sum of the factors that it found to disfavor a finding of fair use. Thus, a stampede score of -4 indicates that the opinion found all four factors to disfavor fair use, while a score of, say, -2 indicates that the opinion found one factor to favor and three factors to disfavor fair use. Figure 6 makes clear that, regardless of posture, judges do not generally stampede the factors when they find fair use. For example, judges found that all four factors favored a finding of fair use in one out of fifteen preliminary injunction opinions and three out of twelve bench trial opinions.

In opinions in which judges found no fair use, however, the data are more open to interpretation. Judges found that all four factors disfavored fair use in 40.0% of the 35 preliminary injunction opinions and 44.0% of the 25 bench trial opinions that found no fair use. Of the 41 cross-motion opinions that found no fair use, 58.5% found that all four factors favored that result. The problem with these data is that we cannot be sure how many of these opinions addressed frivolous fair use defenses that fully merited a -4 stampede score. Indeed, one index of the frivolousness of the defense, particularly in the cross-motion context, may be the stampede score itself. Yet the data showed no relation between the degree of stampeding in an opinion and, if we accept it as an alternative index of the strength or weakness of the fair use defense, the proportion of that opinion devoted to the defense.¹³⁶ In any event, even on their face, these percentages are not so high as to support the claim that, in practice, the factors “tend to degenerate into post-hoc rationales for antecedent conclusions.”¹³⁷ On the contrary, at least outside of the cross-motion context, the majority of district court opinions that found no fair use were willing to acknowledge, as they were in opinions that found fair use, that at least one factor and sometimes more did not support the overall test outcome.

¹³³ See *Sony Corp. of America v. Universal City Studios, Inc.*, 464 U.S. 417, 419 (1984).

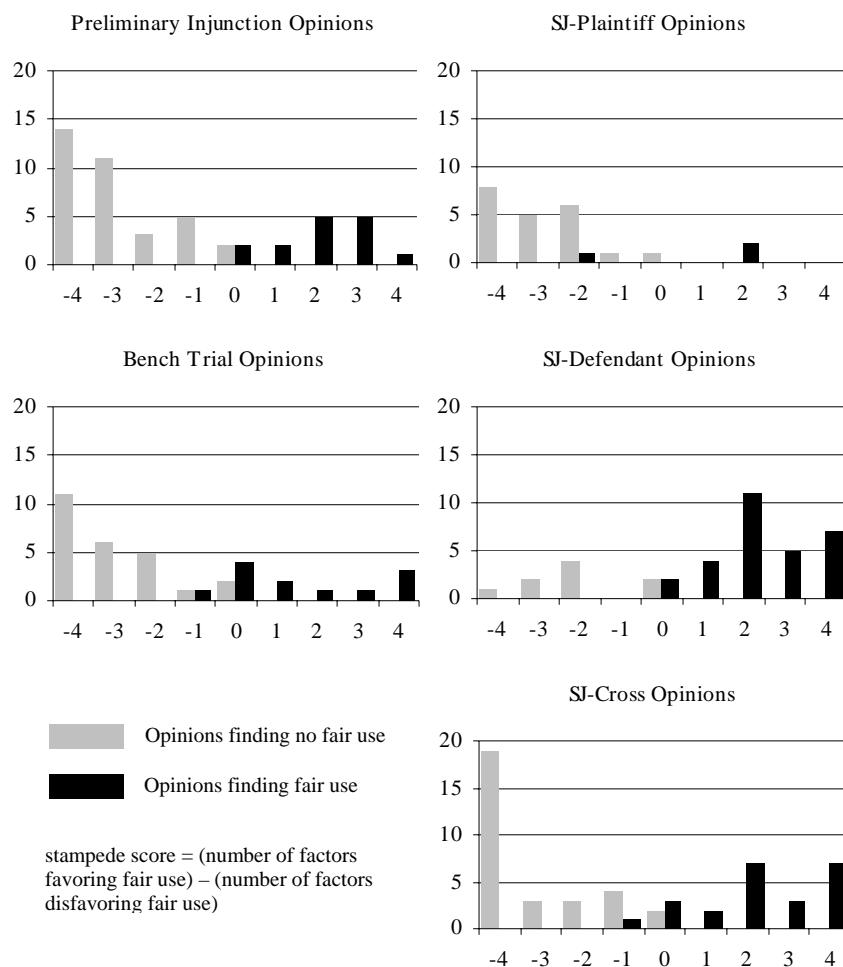
¹³⁴ See *Sony*, 464 U.S. at 457.

¹³⁵ *Harper & Row Publishers, Inc. v. Nation Enterprises*, 471 U.S. 539 (1985) (majority finding all four factors disfavored fair use with dissent finding all four factors favored fair use), *rev’d* 723 F.2d 195 (2d Cir. 1983) (majority finding all four factors favored fair use with dissent finding at least two factors disfavored fair use), *rev’d* 501 F.Supp. 848 (S.D.N.Y. 1980) (finding all four factors disfavored fair use).

¹³⁶ A histogram like that appearing in Figure 4 but indicating the mean of the absolute value of the opinions’ stampede scores rather than their fair use win rate would show a flat stampede score line running across the bins.

¹³⁷ 4 DAVID NIMMER, NIMMER ON COPYRIGHT § 13.05 (2005).

FIGURE 6
 DISTRIBUTION OF FOUR-FACTOR STAMPEDE SCORES BY OUTCOME AND POSTURE IN 208 DISTRICT COURT OPINIONS



Three district court opinions were excluded from the sample used to construct this figure. See Chicago Sch. Reform Bd. of Trustees v. Substance, Inc., 79 F. Supp. 2d 919 (N.D. Ill. 2000) (denying motion to dismiss on a stampede score of -4); Int-Elect Eng'g, Inc. v. Clinton Harley Corp., No. 92-20718, 1993 U.S. Dist. LEXIS 11510 (S.D. Cal. 1993) (denying motion to dismiss on a stampede score of -1); Roy Export Co. Establishment etc. v. Columbia Broadcasting System, Inc., 503 F. Supp. 1137 (S.D.N.Y. 1980) (denying defendant's motion for judgment notwithstanding the verdict on a stampede score of -4).

Finally, from the 215 cases that produced the opinions studied, 32 cases produced both a district court opinion and a circuit court opinion that reviewed the fair use reasoning of the district court opinion. For these 32 cases, Table 8 crosstabulates the stampede score of the district court opinion with the stampede score of the reviewing circuit court decision. In the fifteen cases in which the circuit court reversed the district court on the fair use issue, the difference between the two courts' stampede scores was in certain instances fairly dramatic. Yet in eleven of these fifteen cases, neither the district nor the circuit court opinions yielded stampede scores with absolute values of three or four. In other words, in these eleven cases, both the district and the circuit courts, even when reversing, declined fully or even partially to conform their

TABLE 8
CROSSTABULATION OF THE STAMPEDE SCORES OF APPEALED DISTRICT
COURT OPINIONS BY THE STAMPEDE SCORES OF THE REVIEWING CIRCUIT
COURT DECISIONS, IN 32 FAIR USE CASES

		Stampede Score of Reviewing Circuit Court Opinion									
		-4	-3	-2	-1	0	1	2	3	4	
		-4	4	1		1			1	1	
		-3			1			1	1	1	
		-2	1		2	<u>1</u> [*]					
		-1			<u>1</u> [†]						
Stampede Score of Appealed District Court Opinion		0	<u>1</u>				1				
		1									
		2	<u>1</u>		<u>2</u>			2			
		3	<u>1</u>	<u>1</u>	<u>1</u>						
		4			<u>1</u>			1	1	2	

The number in each box of the crosstabulation indicates the number of cases producing a district court opinion and an appellate court majority opinion meeting the values of the x and y coordinates of the crosstabulation. For example, four cases produced district court and appellate court majority opinions that both yielded stampede scores of -4. Underlined numbers indicate the number of cases in which the appellate court reversed the district court. ^{*}In Association of American Medical Colleges v. Cuomo, 928 F.2d 519 (2d Cir. 1991), the Second Circuit reversed the district court's granting of summary judgment to the plaintiff in Association of American Medical Colleges v. Carey, 728 F. Supp. 873 (S.D.N.Y. 1990). [†]In Veeck v. S. Bldg. Code Congress Int'l, Inc., 241 F.3d 398 (5th Cir. 2001), the Fifth Circuit reversed the district court's granting of summary judgment to the defendant in Veeck v. S. Bldg. Code Congress Int'l, Inc., 49 F. Supp. 2d 885 (E.D. Tex. 1999).

factor analysis to the overall test outcome. Meanwhile, in nine of the seventeen cases in which the circuit court affirmed the district court, the circuit court did not call the factors as the district court had. Here, again, we see no evidence of any strong inclination on the part of the circuit courts to bend the factor outcomes one way or the other. Rather, they appear simply to have called the factors as they saw them.

IV. INTRAFACCTOR ANALYSIS

Having established in previous Parts a macro view of the Section 107 test and the factors that structure it, I seek in this Part to look within the factors themselves to determine which subfactor considerations drive the outcomes of the factors, and through these factor outcomes, the outcome of the overall test. To aid in this inquiry, Table 9 sets out the results of a logistic regression model of the outcome of the fair use test as a function of (1) a variety of factual findings made by judges in the 297 dispositive opinions¹³⁸ and (2) whether the opinion was written by a district or circuit court of the Second or the Ninth Circuits. The results of this model, which correctly classified 85.1% of the 297 opinion outcomes, propound a number of surprising hypotheses that will merit closer investigation in what follows.

Through the course of this primarily descriptive account, the Part will also consider a number of more theoretical or at least generalizable themes. One is what might be termed the irony of the inverted precedent. We will see repeatedly that though the Supreme Court established, for example, that a finding of “X” disfavors fair use, the Court said nothing about a finding of “not X.” The irony emerges in that in the case law that followed, a finding of “X” appeared to have no significant effect on a court’s fair use determination, but a finding of “not X” ended up exerting a significant effect in favor of a determination of fair use. This irony has played out in a number of areas of subfactor doctrine.

A second general theme goes to the concept of “nonergodicity” in systems theory and, a fortiori, in legal precedent.¹³⁹ This concept posits that as a system develops, small, even trivial, initial events may end up having an enormous impact on the system in the long run. In the fair use context, the initial events were Supreme Court dicta that ended up powerfully changing the course of our

¹³⁸ On the problem of the “circularity of facts” in judicial opinions, see Hall & Wright, *supra* note 15.

¹³⁹ See generally Oona A. Hathaway, *Path Dependence in the Law: The Course and Pattern of Legal Change in a Common Law System*, 86 IOWA L. REV. 601 (2001). Cf. Johan Deprez, *Risk, Uncertainty, and Nonergodicity in the Determination of Investment-Backed Expectations: A Post-Keynesian Alternative to Posnerian Doctrine in the Analysis of Regulatory Takings*, 34 LOY. L.A. L. REV. 1221 (2001) (applying economic theories of ergodicity and nonergodicity to regulatory takings doctrine).

TABLE 9
LOGISTIC REGRESSION OF THE OUTCOME OF THE FAIR USE TEST AS A FUNCTION OF FACTUAL FINDINGS
IN 297 DISPOSITIVE OPINIONS

Dependent variable: Fair Use Found (1) / Fair Use Not Found (0)						N=297	Pseudo R ² =0.425
						Correctly Classified: 85.14%	
	Factual Finding	Odds Ratio	Coefficient	Standard Error	P> Z	95% C.I.	
Factor One	D's use is for commercial purpose	.702	-.353	.403	.381	-1.143	.437
	D's use is for non-commercial use	9.466*	2.248	.580	.000	1.111	3.385
	D's use is transformative use	33.572*	3.514	.843	.000	1.861	5.167
	D's use is for parodic purpose	11.215*	2.417	.930	.009	.594	4.241
	D's use is for educational purpose	.377	-.976	.701	.164	-2.349	0.398
	D's use is for research purpose	1.263	.233	.680	.731	-1.099	1.566
	D's use is for critical purpose	1.369	.314	.681	.645	-1.020	1.648
	D accessed P's work improperly	3.761	1.325	.779	.089	-.202	2.851
	D's use is bad faith use	.116	-2.150	1.156	.063	4.416	.115
Factor Two	P's work is a creative work	.402*	-.912	.406	.025	-1.707	-.117
	P's work is a factual work	2.979*	1.091	.484	.024	1.42	2.041
	P's work is unpublished	1.031*	.031	.569	.957	-1.084	1.146
	P's work is published	3.631*	1.290	.611	.035	.092	2.487
Factor Three	D took entirety of P's work	.247*	-1.399	.409	.001	-2.201	-.597
	D took heart of P's work	.018*	-4.019	.945	.000	-5.870	-2.167
	Litigated in Second Circuit	2.348*	.853	.428	.046	.014	1.693
	Litigated in Ninth Circuit	1.993	.690	.442	.119	-.177	1.557
	constant		-.622	.385	.107	-1.377	.134

* denotes statistical significance at the .05 level. The "Odds Ratio" indicates, for example, that when a court finds that the defendant's use was for a non-commercial purpose, the odds of the court finding fair use increase by a factor of 9.5.

fair use doctrine.¹⁴⁰ Related to this is the theme of the perdurability of overturned legal precedent. Even after the Supreme Court sought to correct its initial mistakes, lower courts continued to cite to those mistakes, particularly when they supported the outcome that the lower court reached.

Finally, and related in turn to the perdurability of overturned precedent, this Part will consider the operation of “syntactic” versus “cybernetic feedback”¹⁴¹ in the context of accumulating precedent.¹⁴² This terminology, taken from communications theory, is infelicitous, but it is very important to understanding why our fair use doctrine has to some extent run off the rails of Section 107. The Supreme Court has sought to correct its mistakes in its fair use case law primarily by means of “syntactic feedback.” Rather than admit that its initial communication was inapposite to the world it meant to describe (cybernetic feedback), the Court has invariably stated that lower courts have simply misunderstood its initial communication (syntactic feedback).¹⁴³ In other words, the Court has repeatedly sought to reconstrue what it should have explicitly rescinded and replaced. This practice has proven to be a disaster for our fair use doctrine.

A. Factor One: Purpose and Character of the Use

Factor one calls upon courts to consider “the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit

¹⁴⁰ See, e.g., *id.* at 629 (“[T]he increasing returns nature of the common law system has three predictable consequences. The first is nonergodicity—small early events have a large impact on the eventual outcome. Because judges cannot fully anticipate the ways in which public and private actors will react to their decisions, they have difficulty determining the likely effect of their decisions over time. For this reason, decisions may have large, unanticipated, and unintended effects. For example, a single sentence in an early case may take on increasing significance over time.”).

¹⁴¹ See Martin Shapiro, *Towards a Theory of “Stare Decisis,”* J. LEGAL STUD. 125 (1972). Shapiro’s work deserves far more attention than it has received. Shapiro differentiates between syntactic and cybernetic feedback as follows:

It is important to distinguish syntactic from cybernetic feedback. The former involves transmission back concerning error in the sense of incorrect transmission or receipt of information between sender and receiver within the system; the latter involves transmission concerning error in the sense of incorrect adjustment by the system to the outside world. Thus high levels of syntactic feedback indicate trouble in the transmission facilities of the system, rather than the sensitivity and learning that are typically imputed where high levels of cybernetic feedback are present.

Id. at 126.

¹⁴² On accumulating precedent, see Stefanie Lindquist & Frank B. Cross, *Empirically Testing Dworkin’s Chain Novel Theory: Studying the Path of Precedent*, 80 N.Y.U. L. REV. 1156 (2005).

¹⁴³ See Shapiro, *supra* note 141, at 133-34. Shapiro explains:

Following the rules of *stare decisis*, requests for legal changes, which are actually inspired by the failure of law to adjust correctly to the environment, and are thus cybernetic feedback, are put in the form of syntactic feedback, statements that some judge or lawyer has not correctly received the real message that was transmitted by the previous cases (their “true principles”). In this way much cybernetic feedback information can be squeezed into a communications system that demands very high levels of redundancy, and it can be squeezed in without interfering with that sense of mutual support necessary to the coordination of non-hierarchical organizations.

Id.

educational purposes.”¹⁴⁴ We saw above that outcomes under this factor correlated very strongly with the outcome of the overall fair use test. Indeed, 95.3% of the 148 opinions that found that factor one disfavored fair use eventually found no fair use, while 90.2% of the opinions that found that the factor favored fair use eventually found fair use. Only factor four can boast of higher percentages. To determine what motivated courts to declare that factor one favored or disfavored fair use, we need to look beneath the statutory language, which appears to have had only limited influence on the actual application of the factor, and evaluate how judges considered factor one’s various subfactors. These are the commercial or non-commercial character of the defendant’s use, the degree to which the defendant’s use is transformative of the plaintiff’s work or otherwise “productive,” the propriety or good faith of the defendant’s conduct, and whether the purpose of the defendant’s use falls within one of the categories of purposes mentioned in the preamble of Section 107. I review here each of these in turn.

1. The Commerciality Inquiry

While the concept of transformativeness has received far more attention in the scholarly commentary,¹⁴⁵ I begin with commerciality because it has received far more attention in the case law, particularly among the lower courts. Of the 306 opinions, 84.0% explicitly considered whether the use was commercial or non-commercial in nature under factor one, while only 38.2% explicitly considered the transformativeness of the defendant’s use under the factor. Furthermore, as Figure 7 shows, courts’ greater attention to commerciality was consistent across time. This is regrettable. It is also inconsistent with the actual statutory language of Section 107 and the expectations of its drafters. Indeed, many commentators and some courts, including the Supreme Court in *Campbell*, have been highly critical, even dismissive, of the commerciality inquiry, primarily on the ground that nearly all expression in our culture is produced for profit or is otherwise income-producing in some sense.¹⁴⁶ In his opinion for the majority in *Campbell*, Justice Souter was moved to quote Samuel Johnson, that “no man but a blockhead ever wrote, except for money.”¹⁴⁷ Yet the commerciality inquiry survives, as the data clearly show.

The perdurability of the commerciality inquiry appears largely to be the result of a series of mistakes by the Supreme Court that began with its 1984

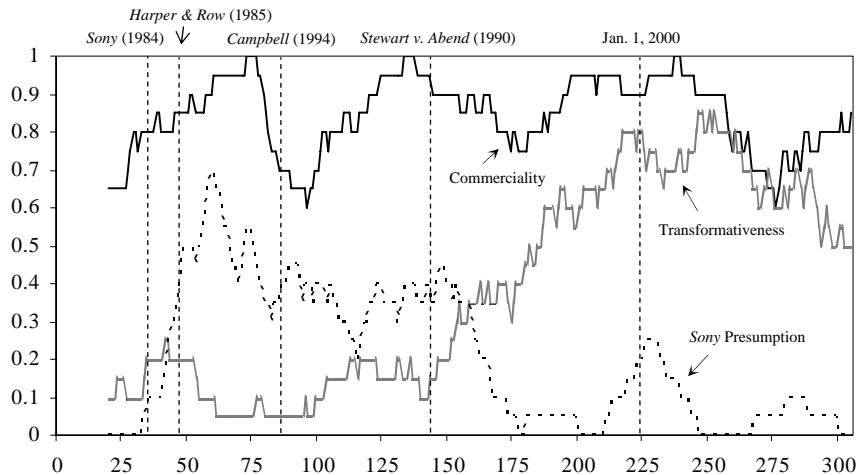
¹⁴⁴ 17 U.S.C. § 107.

¹⁴⁵ See, e.g., Kudon, *supra* note 117; Laura G. Lape, *Transforming Fair Use: The Productive Use Factor in Fair Use Doctrine*, 58 ALB. L. REV. 677 (1995); Diane Leenheer Zimmerman, *The More Things Change the Less They Seem “Transformed:” Some Reflections on Fair Use*, 46 J. COPYRIGHT SOC’Y U.S.A. 251 (1998).

¹⁴⁶ See especially Zimmerman, *supra* note 145.

¹⁴⁷ *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 584 (1994) (quoting 3 BOSWELL’S LIFE OF JOHNSON 19 (G. Hill ed. 1934)).

FIGURE 7
20-OPIINION MOVING AVERAGE OF THE PROPORTION OF OPINIONS MAKING THE COMMERCIALITY AND TRANSFORMATIVENESS INQUIRIES AND CITING THE *Sony* COMMERCIAL USE PRESUMPTION



Sony opinion, mistakes from which fair use doctrine, at least as it is promulgated in our fair use cases, has yet fully to recover. Writing for the majority in *Sony*, Justice Stevens set forth near the end of his opinion what came to be called the “*Sony* presumption.”¹⁴⁸ He did so while discussing under factor four the effect of home video recording on the value of and market for copyrighted television programs. His statement of the *Sony* presumption, with my emphases added, has caused so much trouble in the U.S. fair use case law that it deserves to be quoted in full:

Thus, although *every commercial use of copyrighted material is presumptively an unfair exploitation of the monopoly privilege* that belongs to the owner of the copyright, noncommercial uses are a different matter. A challenge to a noncommercial use of a copyrighted work requires proof either that the particular use is harmful, or that if it should become widespread, it would adversely affect the potential market for the copyrighted work. Actual present harm need not be shown; such a requirement would leave the copyright holder with no defense against predictable damage. Nor is it necessary to show with certainty that future harm will result. What is necessary is a showing by a preponderance of the evidence that some meaningful likelihood of future harm exists. *If the intended use is for commercial gain, that*

¹⁴⁸ See, e.g., James Boyle, *Intellectual Property Policy Online: A Young Person's Guide*, 10 HARV. J.L. & TECH. 47, 99 (1996); Stacey L. Dogan, *Comment: Sony, Fair Use, and File Sharing*, 55 CASE W. RES. L. REV. 971, 973 (2005).

likelihood may be presumed. But if it is for a noncommercial purpose, the likelihood must be demonstrated.¹⁴⁹

This was a quite stunning interpretation of Section 107, especially since the first factor referred to the “commercial nature” of the defendant’s use as only one consideration among others that the factor encompassed, and even more so since the dependent clause in which the reference appeared was inserted very much at the eleventh hour of the drafting process, primarily to address the concerns of those who were engaged in “nonprofit educational purposes.”¹⁵⁰ The next year in *Harper & Row*, the Court only made matters worse. Clearly sensing that it had overreached, the Court sought to bend its previous statement: “The fact that a publication was commercial as opposed to nonprofit is a separate factor that tends to weigh against a finding of fair use.”¹⁵¹ So far, so good. But then the Court quoted the very dictum from *Sony* that it was seeking to overwrite—“every commercial use of copyrighted material is presumptively an unfair exploitation of the monopoly privilege”¹⁵²—as if a tendency to disfavor fair use and a presumption against fair use were the same. In an effort to maintain the appearance of “redundancy,”¹⁵³ the Court resorted to “syntactic feedback” where “cybernetic feedback” was in order.¹⁵⁴ Then, in its 1990 *Stewart* opinion, the Court again quoted the same phrase from *Sony* that *Harper & Row* had, but this time *without* the *Harper & Row* modification.¹⁵⁵ Finally, in 1994, the *Campbell* Court made it clear that the *Sony* presumption was no longer good law. It did so, however, syntactically, through reconstrual: “[A]s we explained in *Harper & Row*, *Sony* stands for the proposition that the ‘fact that a publication was commercial as opposed to nonprofit is a separate factor that tends to weigh against a finding of fair use.’”¹⁵⁶

And so, through *Sony*, the *Harper & Row* gloss on *Sony*, *Stewart*’s revival of *Sony*, and then *Campbell*’s revival of the *Harper & Row* gloss, we see not so much a refinement of precedent as an accumulation of precedent. The sheer mass of this precedent, perhaps even regardless of what it said, appears to have kept the commerciality inquiry in the foreground of the fair use analysis far from where the drafters of Section 107 originally intended that it should be. Figure 7 suggests that courts’ attention to the commerciality inquiry peaked soon after *Sony* and *Harper & Row* and then declined until *Stewart*, which appears to have revived it. The proportion of courts conducting the

¹⁴⁹ *Sony Corp. of America v. Universal City Studios, Inc.*, 464 U.S. 417, 451. (1984)

¹⁵⁰ See PATRY, *supra* note 28, at 351-53.

¹⁵¹ *Harper & Row Publishers, Inc. v. Nation Enterprises*, 471 U.S. 539, 562 (1985).

¹⁵² *Id.* (quoting *Sony*, 464 U.S. at 451).

¹⁵³ See Shapiro, *supra* note 141, at 127 (“[L]egal discourse organized by the rules of *stare decisis* emphasizes, and itself insists that its success rests upon, high levels of redundancy.”).

¹⁵⁴ Cf. *id.* at 133 (“If the system employs high levels of syntactic redundancy, it does not have the ‘space’ to transmit much cybernetic feedback information to its receiving parts.”).

¹⁵⁵ *Stewart v. Abend*, 495 U.S. 207, 237 (1990).

¹⁵⁶ *Campbell*, 510 U.S. at 585.

commerciality inquiry declined somewhat after *Campbell* but then returned more or less to its pre-*Campbell* level.

What about courts' treatment of the *Sony* presumption in particular? Figure 7 suggests that the proportion of opinions citing the *Sony* presumption peaked soon after *Sony* and *Harper & Row*, and then declined to a level that held until *Campbell*, the Court opinion which explicitly, albeit syntactically, repudiated the presumption and should have buried it once and for all. Writing soon after the Supreme Court handed down *Campbell*, Judge Leval spoke of it as "Justice Souter's rescue of fair use."¹⁵⁷ Before *Campbell*, Judge Leval explained, courts would invoke the *Sony* presumption if they otherwise thought that the commercial use at issue was unfair, but if they wanted instead to find fair use, they "would simply omit any mention of the 'commercial' thing."¹⁵⁸ Now, *Campbell* "has dispelled all those unhelpful slogans from the fair use discussions—particularly the pernicious 'commercial use' presumption."¹⁵⁹

The data do not bear this out. On the contrary, they suggest that Judge Leval may have been overly pessimistic with respect to how judges used the *Sony* presumption before *Campbell*, but overly optimistic with respect to how they would use it after *Campbell*. Of the 108 opinions produced between *Sony* and *Campbell*, 45 (41.7% of 108) cited the *Sony* presumption. Of these, 29 found the defendant's use to be commercial and 14 found fair use, for a fair use win rate in these opinions of .311. Of the 61 opinions during that period that did not cite the *Sony* presumption, 29 found fair use, for a fair use win rate of .475. Though statistically significant, the difference in win rates between the two sets of opinions does not, of course, support the proposition that courts invoked the *Sony* presumption only when finding for the plaintiff. For the 162 opinions produced after *Campbell*, however, the story is somewhat different, and contrary to Judge Leval's hopes. As Figure 7 shows, *Campbell* triggered a decline in the proportion per year of opinions citing the *Sony* presumption, but this decline was then followed by a renewal of interest in the presumption among some lower courts. Specifically, 12 of the 162 opinions produced after *Campbell* invoked the *Sony* presumption, with eleven of these twelve finding that the defendant's use was commercial, that factor one favored the plaintiff, and ultimately that there was no fair use.¹⁶⁰ This is strongly suggestive of certain courts' cynical—or at best, unknowing—use of the *Sony* presumption notwithstanding *Campbell*.

Overall, despite the language of Section 107, the commerciality inquiry and the *Sony* presumption in particular remain exceptionally tenacious memes in the fair use case law. No doubt this reflects in part their high fitness for a litigation

¹⁵⁷ Pierre N. Leval, *Campbell v. Acuff-Rose: Justice Souter's Rescue of Fair Use*, 13 CARDozo ARTS & ENT. L.J. 19 (1994).

¹⁵⁸ *Id.* at 20-21.

¹⁵⁹ *Id.* at 22.

¹⁶⁰ See, e.g., *Itar-Tass Russian News Agency v. Russian Kurier*, 886 F. Supp. 1120 (S.D.N.Y. 1995).

environment pervaded with commercial expression. But it is also a consequence of the Supreme Court’s repeated attempts to maintain appearances by reconstruing what it should simply have rescinded and replaced. Even so, there is a strange final irony to all of this attention to the commerciality of the defendant’s use. The regression results presented in Table 9 suggest that notwithstanding the *Sony* presumption, and notwithstanding *Campbell*’s revival of the *Harper & Row* gloss that a “commercial as opposed to [a] nonprofit [use] is a separate factor that tends to weigh against a finding of fair use,”¹⁶¹ a finding that the defendant’s use was for a commercial purpose (which was made in 64.4% of the opinions) did not significantly influence the outcome of the fair use test in favor of an overall finding of no fair use. Rather, it was a finding that the defendant’s use was for a *non-commercial* purpose (which was made in 15.4% of the opinions) that strongly influenced the outcome of the test in favor of an overall finding of fair use. We are familiar with vertical and horizontal precedent. But if there is such a thing as empirical or popular precedent, based on the empirical analysis of how the population of judges who have previously employed a legal doctrine did so, then this is one such precedent worth following. While the fact that a defendant’s use is for a commercial purpose should generally play no significant role in a fair use determination, the fact that a defendant’s use is for a non-commercial purpose should be understood, as it appears it already is in practice, strongly to support a finding of fair use.

2. The Transformativeness Inquiry

In his 1990 Harvard Law Review article *Towards a Fair Use Standard*,¹⁶² Judge Leval encouraged courts to attend more closely in their fair use analyses to the degree to which a defendant’s use was “transformative.” Judge Leval argued that in weighing “the strength of the secondary user’s justification against factors favoring the copyright owner,” the court should consider if “the secondary user adds value to the original—if the quoted matter is used as raw material, transformed in the creation of new information, new aesthetics, new insights and understandings,” because “this is the very type of activity that the fair use doctrine intends to protect for the enrichment of society.”¹⁶³ Four years later, the *Campbell* court relied heavily on the concept of transformativeness and on Judge Leval’s exposition of it in its factor one analysis of 2 Live Crew’s parody.¹⁶⁴ Courts and commentators have since spoken of the concept of transformativeness as the cynosure of fair use analysis: we are told that it is “vitally important to the fair use inquiry,”¹⁶⁵ it

¹⁶¹ *Campbell*, 510 U.S. at 585.

¹⁶² Leval, *supra* note 92.

¹⁶³ *Id.* at 1111.

¹⁶⁴ *Campbell*, 510 U.S. at 578-79.

¹⁶⁵ Leval, *supra* note 92, at 1111.

goes to the “heart of the fair use inquiry.”¹⁶⁶ As one court has put it, “[a]lthough ‘transformativeness’ is primarily analyzed in connection with the first fair use factor, it forms the basis of the entire fair use analysis.”¹⁶⁷ Indeed, some have suggested that the transformativeness inquiry has essentially superseded Section 107 as the backbone of our fair use doctrine.¹⁶⁸

It appears, however, that courts and commentators have exaggerated the influence of transformativeness doctrine on our fair use case law. At the district court level, 41.2% of the 119 district court opinions following *Campbell* failed even to refer to the doctrine, while 90.2% of the 92 opinions preceding *Campbell* failed to reference it (either under the rubric of transformativeness or, as it was sometimes called, “productive use”¹⁶⁹). At the circuit court level, the percentages are better, but still far from what we would expect for a doctrine said now to be the basis of the fair use inquiry. Of the 43 circuit court opinions that followed *Campbell*, 18.6% failed to invoke the concept, while 84.4% of the 45 circuit court opinions that preceded *Campbell* failed to invoke it. Furthermore, the doctrine appears to be losing strength. Figure 7 shows the proportion of opinions over time that made some reference to transformativeness. The doctrine’s citation curve began its downward slope sometime in the early part of this decade.

Nevertheless, in those opinions in which transformativeness did play a role, it exerted nearly dispositive force not simply on the outcome of factor one but on the overall outcome of the fair use test. More specifically, the data suggest that while a finding of transformativeness is not necessary to trigger an overall finding of fair use, it is sufficient to do so. We can infer that it is not necessary in light of the fact that 25 (or 36.8%) of the 68 post-*Campbell* opinions that found fair use made no reference to transformativeness and four explicitly found that the defendant’s use was not transformative. We can infer that it is sufficient (or nearly sufficient) in light of the fact that each of the 13 circuit court opinions and 27 of the 29 district court opinions that found the defendant’s use to be transformative also found it to be a fair use—and one of

¹⁶⁶ On *Davis v. Gap, Inc.*, 246 F.3d 152, 171 (2d Cir. 2001) (Leval, J.).

¹⁶⁷ *Hofheinz v. Discovery Communications, Inc.*, No. 00-3802, 2001 WL 1111970, at *3 (S.D.N.Y. 2001).

¹⁶⁸ Cf. 4 NIMMER, *supra* note 137, at § 13.05 (“Those Second Circuit cases appear to label a use ‘not transformative’ as a shorthand for ‘not fair,’ and correlatively ‘transformative’ for ‘fair.’ Such a strategy empties the term of meaning—for the ‘transformative’ moniker to guide, rather than follow, the fair use analysis, it must amount to more than a conclusory label.”); Sag, *supra* note 11, at 388 (“[T]ransformativeness is clearly a meta-factor: the extent to which a use transforms the work cannot be determined without reference to the other factors, such as the nature of the original work, the quantitative and qualitative similarity between the works and the effect of the use on the value of the original work.”); *Id.* (“The dominance of the transformativeness test makes the actual statutory language regarding non-commercial and educational uses largely irrelevant.”).

¹⁶⁹ See, e.g., *Universal City Studios, Inc. v. Sony Corp. of America*, 659 F.2d 963, 970 (9th Cir. 1981) (“As the first sentence of § 107 indicates, fair use has traditionally involved what might be termed the ‘productive use’ of copyrighted material.”).

the two district court outliers was reversed on appeal.¹⁷⁰ Furthermore, as several courts explicitly noted,¹⁷¹ a finding of transformativeness trumped a finding that the defendant's use was commercial for purposes of determining whether factor one favored fair use. In 28 opinions, the court found the defendant's use to be both commercial and transformative under factor one, and in 26 of these opinions, the court found both that factor one and the overall test favored fair use—with one of the two outliers reversed on appeal.

These data may go far towards explaining why Table 9 reports such a high coefficient and odds ratio for a finding of transformativeness. Putting the table's regression results in more easily understood terms, consider that, on the regression model used to produce the table, a defendant has a 35.5% chance of winning the fair use defense where it has made a non-transformative, commercial use of a creative, published work (with all other variables set at zero). If that same use were found to be transformative, the defendant's chance of winning the fair use defense would increase to 94.9%.

Interestingly, while a finding of transformativeness may be dispositive of the outcome of the fair use test, such a finding does not stampede the outcomes of the other factors. Of the 43 opinions that found the defendant's use to be transformative, only six found that all four factors favored a finding of fair use. Instead, 26 of these 43 opinions found that factor two, going to the nature of the plaintiff's work, disfavored fair use, and 6 of these 26 further found that factor three, going to the amount and substantiality of the defendant's taking, also disfavored a finding of fair use. This makes sense. Defendant's are far more likely to make a transformative use of a creative rather than a factual work, and their transformative use is likely to involve a substantial taking of the plaintiff's expression. To their credit, in opinions addressing a transformative use, courts were generally willing to call factors two and three as they saw them rather than bend those factor outcomes to conform to the overall test outcome. There were exceptions, however. Among those opinions that found that defendant's use was transformative, but that factors two or three nevertheless favored fair use, were several fine examples of judicial slight of hand.¹⁷²

¹⁷⁰ See *SunTrust Bank v. Houghton Mifflin Co.*, 136 F. Supp. 2d 1357 (N.D. Ga. 2001), vacated, 252 F.3d 1165 (11th Cir. 2001) and 268 F.3d 1257 (11th Cir. 2001). The one outlier was *Castle Rock Entertainment v. Carol Publ. Group*, 955 F. Supp.260 (S.D.N.Y. 1997), *aff'd*, 150 F.3d 132 (2d Cir. 1998).

¹⁷¹ See, e.g., *Leibovitz v. Paramount Pictures Corp.*, 948 F. Supp. 1214, 1223 (S.D.N.Y. 1996) ("I find that the purposes of copyright are best served by a finding that the highly transformative character of the Nielsen ad trumps its admittedly commercial purpose and that the first fair use factor therefore weighs in favor of the defendant, albeit perhaps by only a slight margin."); *American Geophysical Union v. Texaco, Inc.*, 802 F. Supp. 1, 12-13 (S.D.N.Y. 1992) ("Thus courts have repeatedly found in favor of transformative secondary uses on the first factor, notwithstanding the presence of profit motivation. Thus, although courts ritually proclaim, almost as a mantra, that every commercial use is 'presumptively' unfair, that presumption is easily overcome by a transformative, nonsuperseding use." (citations omitted)).

¹⁷² See, e.g., *Nunez v. Caribbean Int'l News Corp.*, 235 F.3d 18, 24 (1st Cir. 2000) (stating under the third factor that "[i]n this case, El Vocero admittedly copied the entire picture; however, to copy any less than that would have made the picture useless to the story. As a result, like the district court, we count this factor as of little consequence to our analysis."); *Blanch v. Koons*, 396 F. Supp. 2d 476, 481-82 (S.D.N.Y. 2005) ("The

3. The Bad Faith Inquiry

Must a defendant's use be "fair" in order to qualify as a "fair use"? Should the "propriety" of the defendant's conduct play any role in the court's fair use analysis? Lloyd Weinreb thought so, and urged courts to consider the "fairness" of the defendant's use as an additional overarching factor.¹⁷³ Other commentators have questioned the feasibility of the fairness inquiry on the ground that fairness is too subjective—or circular—a standard.¹⁷⁴ Still others argue that the fairness inquiry unnecessarily complicates the fundamental utilitarian question, going to what outcome will ultimately promote the progress of human creativity, that should inform any fair use determination.¹⁷⁵

The data suggest that considerations of fairness, propriety, or good or bad faith have not played a significant role in our fair use case law—this notwithstanding the frequency with which the opinions intoned that fair use is an "equitable doctrine."¹⁷⁶ However, in the few cases in which courts explicitly found that the defendant's conduct was undertaken in bad faith, courts almost invariably found no fair use. Forty-nine (or 16.0%) of the 306 opinions made some reference to the propriety of the defendant's conduct or to the relevance of the fairness analysis, with 25 of them doing so under factor one.¹⁷⁷ Fourteen opinions found improper conduct and twelve of these found no fair use, with seven finding that all four factors favored that overall outcome. Conversely, of

photograph as a whole is sufficiently creative and original to receive copyright protection, and it is copyrighted. Yet its general publication throughout the United States favors fair use . . . Blanch has no right to the appearance of the Gucci sandals (perhaps the most striking element of the photograph), and Koons appropriated nothing else of the photograph except the crossed legs. Viewed alone (disregarding the sandals) they are banal rather than creative. The second factor favors the defendants." (citations omitted); Newport-Mesa Unified Sch. Dist. v. Cal. Dep't of Educ., 371 F.Supp.2d 1170, 1177-78 (C.D. Cal. 2005) ("Under the second fair use factor, the nature of the copyrighted work is creative rather than informational. Development of the test questions as well as their compilation in a particular test form is a creative, imaginative, and original process. This ordinarily would weigh against finding fair use. But, with the addition of a student's answers, the questions and answers are informational in nature, which weighs in favor of fair use." (citations omitted)); Penelope v. Brown, 792 F.Supp. 132, 138 n. 7 (D. Mass. 1992) (stating under the second factor that "Penelope's work is scholarly and hence that the law should favor its dissemination").

¹⁷³ See Weinreb, *supra* note 19, at 1138 ("[F]air use has historically been and ought to remain what its name suggests: an exemption from copyright infringement for uses that are fair."); *id.* at 1141 ("Although the courts were presumably construing the statute according to the legislative intent, it has from the beginning had the flavor of an equitable doctrine, importing, as its name indicates, considerations of fairness not directly related to the statutory purpose.").

¹⁷⁴ See, e.g., BRUCE P. KELLER & JEFFREY P. CUNARD, COPYRIGHT LAW § 8:1 ("The central concept underlying the doctrine—an assessment of the 'fairness' of the use in question, based on a balancing of several factors—is inherently subjective. As a result, what one judge in his or her personal view regards as a 'fair use' may sharply contrast with what another judge down the hall may think.").

¹⁷⁵ See, e.g., Leval, *supra* note 92, at 1125-26; Madison, *supra* note 9, at 1555-56.

¹⁷⁶ By this, courts typically meant simply that that fair use doctrine is "an equitable rule of reason[] which permits courts to avoid rigid application of the copyright statute when, on occasion, it would stifle the very creativity which that law is designed to foster." *Stewart v. Abend*, 495 U.S. 207, 236 (1990) (citations omitted).

¹⁷⁷ One opinion considered the defendant's bad faith under both the first factor and as an additional factor. *See Haberman v. Hustler Magazine, Inc.*, 626 F. Supp. 201, 211 (D. Mass. 1986) (under factor one); *id.* at 214 (as additional factor).

the 28 opinions in which the court explicitly found the defendant's conduct not to be improper, 24 found fair use. Unlike in their opinions finding bad faith, courts did not tend to stampede the factors in opinions finding good faith.¹⁷⁸

Though they strongly suggest that a finding of bad faith will trigger a finding of no fair use, the data should be interpreted with caution. More so than any other subfactor consideration, the fairness determination appears to be susceptible to a basic circularity: did the court find the defendant's use to be unfair because the defendant's conduct was improper, or did the court find the defendant's conduct to be improper because its use was unfair? The opinions that made a fairness determination tended to do so on highly fact-specific grounds, so no strong generalizations can be made to help answer this question. Still, there was only one opinion in which the court explicitly found that but for the defendant's bad faith, the court would have found fair use.¹⁷⁹ Otherwise, the fairness consideration appeared rarely to be rarely decisive.¹⁸⁰ This is consistent with the regression results reported in Table 9, in which the bad faith variable failed to produce a significant coefficient. More commonly, a fairness finding appeared to function as merely an additional consideration in support of an outcome already determined—or overdetermined—by other considerations.

4. The Preambular Purposes Inquiry

Section 107 volunteers in its preamble certain examples of fair purposes. Specifically, it provides that the defendant's use of the plaintiff's expression might be deemed fair if it is made “for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research.”¹⁸¹ Table 10 sets out the number and proportion of opinions which found the defendant's use to fall within one of these categories. While defendants generally enjoyed very high fair use win rates in these opinions, defendants engaged in “educational” purposes did not. There is nothing remarkable about the 22 educational purpose opinions that can explain this surprising result.¹⁸² In any event, the regression model suggests that, when controlling for the effects of other findings, a finding that the defendant's use

¹⁷⁸ The mean stampede score in the 14 opinions that found bad faith was -2.43 (SD=2.06), while the mean stampede score in the 28 opinions that found no bad faith was 1.82 (SD=2.31).

¹⁷⁹ See *Softel, Inc. v. Dragon Medical & Scientific Communs.*, No. 87-0167, 1992 U.S. Dist. LEXIS 9502 (S.D.N.Y. 1992) (“In this case, defendants' lack of good faith in utilizing the routines tips the balance of the factors against a finding of fair use. Accordingly, plaintiff is entitled to recover on its claim of copyright infringement with respect to defendants' literal copying of plaintiff's image retrieval routines in the programs Hairy Cell Roche and Low Back Pain.”). *Softel* yielded a stampede score of 0.

¹⁸⁰ See, e.g., *New Line Cinema Corp. v. Bertlesman Music Group, Inc.*, 693 F. Supp. 1517, 1530 (S.D.N.Y. 1988) (“The Court's conclusion is also bolstered by examining the propriety of Zomba's conduct in this case.”). See also *Rogers v. Koons*, 960 F.2d 301, 309 (S.D.N.Y. 1992); *Original Appalachian Artworks, Inc. v. Topps Chewing Gum, Inc.*, 642 F. Supp. 1031, 1036 (N.D. Ga. 1986).

¹⁸¹ See generally *Madison*, *supra* note 9, at 1552-57.

¹⁸² Only four of these opinions involved standardized testing.

fell within one of the preambular categories did not significantly effect the outcome of the fair use test. Instead, as with a finding of bad faith, other considerations appear to have determined the outcome otherwise endorsed by the preambular purposes inquiry.

TABLE 10
DISTRIBUTION AND FAIR USE WIN RATES OF OPINIONS EXPLICITLY
ADDRESSING PREAMBULAR PURPOSES

Preambular Purpose	N	% of 306	Found FU
Research purpose	22	7.2	.409
Critical purpose	29	9.5	.621
News reporting	27	8.8	.778
Educational purpose	27	8.8	.482

B. Factor Two: Nature of the Copyrighted Work

Factor two instructs courts to consider “the nature of the copyrighted work.”¹⁸³ The data with respect to factor two are seemingly as ambiguous and open to interpretation as the statutory language of the factor itself. Despite Section 107’s command that “the factors to be considered shall include,”¹⁸⁴ factor two, 17.7% of the 306 opinions failed even to refer to the factor, while an additional 6.5% did so only to call it irrelevant. We saw above that the correlation coefficients in Table 5 and the regression coefficients in Table 6 suggest that the outcome of factor two typically has no significant effect on the overall outcome of the fair use test. Many courts and commentators have long asserted as much.¹⁸⁵ Yet the regression coefficients reported in Table 9 for certain subfactor considerations under factor two suggest that we cannot write off the factor entirely. On the contrary, certain findings under factor two appear significantly to affect the outcome of the fair use test, sometimes in ways that run contrary—or obliquely—to the original intent of the doctrine underlying those findings.

From the rather open-ended statutory language of the second factor have emerged two subfactor considerations: whether the plaintiff’s work is creative or factual in nature and whether it is published or unpublished. I consider each of these in turn.

¹⁸³ 17 U.S.C. § 107.

¹⁸⁴ *Id.*

¹⁸⁵ See *supra* note 113.

1. The Creative/Factual Work Inquiry

As the *Campbell* Court explained, factor two “calls for recognition 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.”¹⁸⁶ According to this framework, creative works “of fiction or fantasy”¹⁸⁷ stand at the core of copyright protection, making a finding of their fair use less likely, while factual works stand at the periphery, making a finding of their fair use more likely.

Though courts and commentators have belittled the significance of the creative/factual work inquiry along with the rest of factor two, the data suggest that there was in fact a significant inverse relation between the creativity of the plaintiff’s work and the likelihood of its fair use in the opinions studied. Of the 306 opinions, 41.1% explicitly found that the plaintiff’s work was creative rather than factual in nature, and 34.1% of these opinions found fair use. Meanwhile, 16.3% of the opinions explicitly found that the plaintiff’s work was factual rather than creative in nature, and 54.0% of these opinions found fair use. Though it is probably impossible to establish any degree of causality, it is also interesting to note that opinions in which a court explicitly found the plaintiff’s work to be creative or factual also exhibited a good deal of stampeding. Among the 126 opinions that found the plaintiff’s work to be creative in nature, a strong plurality of 43.4% found all four factors to disfavor fair use, while among the 50 opinions that found the plaintiff’s work to be factual in nature, a plurality of 28.0% found all four factors to favor fair use.¹⁸⁸

It should not be surprising, then, that the regression model reported in Table 9 predicts that a defendant has a 35.5% chance of winning the fair use defense when it makes a commercial use of a published work that is found to be creative in nature (with all other variables set at zero), and an 80.3% of winning the fair use defense when it makes a commercial use of a published work that is found instead to be factual in nature (again, with all other variables set at zero). This is a substantial difference in the probability of success, one that suggests that notwithstanding the correlation and regression coefficients reported in Tables 5 and 6, the creative or factual nature of the plaintiff’s work does in fact play a role in judges’ determination of the fair use issue.

As a normative matter, this again is an empirical precedent altogether worth following. In principle, the four factors of Section 107 seek to “balanc[e] the need to provide individuals with sufficient incentives to create public works

¹⁸⁶ *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 586 (1994).

¹⁸⁷ *Harper & Row Publishers, Inc. v. Nation Enterprises*, 471 U.S. 539, 563 (1985) (“The law generally recognizes a greater need to disseminate factual works than works of fiction or fantasy.”).

¹⁸⁸ An additional 26.0% of these 50 opinions yielded a stamped score of -2, meaning that three factors disfavored and one factor favored fair use. All of these opinions found no fair use, and all but one found that the one factor not supporting that outcome was factor two.

with the public's interest in the dissemination of information.¹⁸⁹ As such, the Section 107 test should do more than simply evaluate, under factor four, the extent to which the defendant's use impacts the "potential market for or value of the copyrighted work." That inquiry addresses only one side of the balance. The first and second factors of Section 107 exist to remind courts to consider the other side of the balance, going to the "public's interest in the dissemination of information."¹⁹⁰ Happily, it appears at least with regard to the creative/factual inquiry that many courts are doing just that.

2. The Published/Unpublished Work Inquiry

The story of the published versus unpublished work inquiry under factor two is a story of nonergodicity, unintended consequences, and accumulating precedent. In her 1984 *Harper & Row* opinion for the majority, Justice O'Connor stated that "[t]he fact that a work is unpublished is a critical element of its 'nature'" and that "the scope of fair use is narrower with respect to unpublished works."¹⁹¹ She also wrote that "[u]nder ordinary circumstances, the author's right to control the first public appearance of his undisseminated expression will outweigh a claim of fair use."¹⁹² Writing in dissent, Justice Brennan declared that this latter statement "introduces into analysis of this case a categorical presumption against prepublication fair use."¹⁹³ Three years later, in *Salinger v. Random House*, the Second Circuit adopted the *Harper & Row* dissent's aggressive interpretation of the majority opinion: "We think that the tenor of the Court's entire discussion of unpublished works conveys the idea that such works normally enjoy complete protection against copying any protected expression."¹⁹⁴ A variety of opinions from the Second Circuit and elsewhere subsequently furthered this reading,¹⁹⁵ so that by 1991, the Second Circuit would declare, in *Wright v. Warner Books*, that "[u]npublished works are the favorite sons of factor two" and "our precedents . . . leave little room for discussion of the factor once it has been determined that the copyrighted work is unpublished."¹⁹⁶ It seemed at the time that the courts had built a new *Sony*-like presumption, this time against uses of unpublished works.¹⁹⁷ In 1992,

¹⁸⁹ *Hustler Magazine Inc. v. Moral Majority Inc.*, 796 F.2d 1148, 1151 (9th Cir. 1986).

¹⁹⁰ See generally Brett M. Frischman & Mark A. Lemley, *Spillovers*, 107 COLUM. L. REV. 257, 286-90 (2007) (discussing the importance of positive externalities, or "spillovers," to a proper understanding of the fair use defense).

¹⁹¹ *Harper & Row Publishers, Inc. v. Nation Enterprises*, 471 U.S. 539, 564 (1985).

¹⁹² *Id.* at 555.

¹⁹³ *Id.* at 595 (Brennan, J., dissenting).

¹⁹⁴ 811 F.2d 90, 97 (2d Cir. 1987).

¹⁹⁵ See, e.g., *New Era Publs. Int'l v. Henry Holt & Co.*, 873 F.2d 576, 583 (2d Cir. 1989) ("Where use is made of materials of an 'unpublished nature,' the second fair use factor has yet to be applied in favor of an infringer, and we do not do so here."); *Association of American Medical Colleges v. Carey*, 728 F. Supp. 873, 885-886 (N.D.N.Y. 1990); *Love v. Kwitny*, 706 F. Supp. 1123, 1134 (S.D.N.Y. 1989).

¹⁹⁶ 953 F.2d 731, 737 (2d Cir. 1991)

¹⁹⁷ See generally Crews, *supra* note 27.

Congress intervened. It amended Section 107 to include at its conclusion: “The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the above factors.”¹⁹⁸

The strange result of this series of events, a result which the *Harper & Row* court could hardly have foreseen, is that in the opinions studied, the fact that the plaintiff’s work was unpublished appears to have exerted no significant effect on the outcome of the fair use test, but the fact that the plaintiff’s work was *published* appears to have exerted a strong effect on the outcome of the test in favor of a finding of fair use (see Table 9). Meanwhile, a wide diversity of precedent developed around the published/unpublished work inquiry, so that courts could draw upon previous case law to assert that both the unpublished or published status of the work either favored or disfavored a finding of fair use. Thirty-seven opinions explicitly found that the plaintiff’s work was unpublished, with 29 of these asserting that this fact disfavored a finding of fair use, three asserting that it favored a finding of fair use (because plaintiff’s expression was not otherwise available to the public),¹⁹⁹ and five concluding that the unpublished status of the plaintiff’s work was unimportant to the fair use analysis. Courts found fair use in roughly half (48.6%) of these opinions. By comparison, 42 opinions explicitly found that the plaintiff’s work was published, with 36 asserting that this fact favored a finding of fair use, four asserting that it disfavored a finding of fair use (because the plaintiff’s work was otherwise available to the purchasing public),²⁰⁰ and two asserting that it made no difference to the outcome of the fair use test. Courts found fair use in 77.8% of the 36 opinions that asserted that the published status of the plaintiff’s work favored a finding of fair use, and in 69.1% of the 42 opinions overall. As with the subfactor variables discussed above, the inner workings of these data help to explain why the regression model predicts (with commercial use set to one and all other variables set to zero) that the defendant’s chances of succeeding in its fair use defense increase from 27.4% to 57.8% when the status of the plaintiff’s work shifts from being not published to published—and not at all when the status of the plaintiff’s work shifts from being unpublished to not unpublished.

From the perspective of a fair use maximalist, the data once again reveal an encouraging, and ironic, result. As above, the Supreme Court sought to establish that a certain finding (here, that the work is unpublished) *disfavors* fair use. Lower courts appear not to have acted on that dictum, however, other

¹⁹⁸ Pub. L. No. 102- 492, 106 Stat. 3145 (1992).

¹⁹⁹ See, e.g., *Maxtone-Graham v. Burtchaell*, 631 F. Supp. 1432, 1437 (S.D.N.Y. 1986) (“The advancement of the social sciences and public discourse on an important issue is enhanced by permitting liberal, but fair, use of such materials.”). Cf. *Penelope v. Brown*, 792 F. Supp. 132, 138 (D. Mass. 1992) (“Although *Teaching About Doublespeak* is still in print and is sold in college bookstores, it cannot be said to be widely available to the public. Lack of availability lends Brown greater justification for reproducing it.”).

²⁰⁰ See, e.g., *Encyclopaedia Britannica Educational Corp. v. Crooks*, 558 F. Supp. 1247, 1251 (W.D.N.Y. 1983); *Encyclopaedia Britannica Educational Corp. v. Crooks*, 542 F. Supp. 1156, 1177 (W.D.N.Y. 1982).

than to invert it to conclude that the opposite of that finding (here, that the work is published) *favors* fair use.²⁰¹

C. Factor Three: Amount and Substantiality of the Use

The third factor instructs courts to consider “the amount and substantiality of the portion used in relation to the copyrighted work as a whole.”²⁰² In other words, and as the great majority of the opinions recognized,²⁰³ this factor calls upon courts to evaluate, both on a quantitative and a qualitative scale, what proportion of the plaintiff’s work the defendant used. We saw above that the outcome of factor three correlated strongly with the outcome of the overall test as well as with the outcomes of factors one and four. These relations were especially strong when the third factor was found to favor fair use. This is not surprising. An insubstantial taking of the plaintiff’s work will not likely have an appreciable effect on the market for or value of that work, and perhaps on that ground alone, will likely trigger a finding of fair use. Indeed, of the 79 opinions that found that factor three favored fair use, 76 subsequently found fair use, and 72 of these also found that factor four favored that result.

Of all of the factors, the third factor boasts the most settled and easily-understood doctrine. In general, the more the defendant takes of the plaintiff’s work, the less likely it is that the taking will qualify as a fair use. What if the defendant takes the entirety of the plaintiff’s work? Courts and commentators have asserted that, “generally, it may not constitute a fair use if the entire work is reproduced,”²⁰⁴ though “there may be certain very limited situations wherein copying of even the entire work for a different functional purpose may be regarded as a fair use.”²⁰⁵ The data tell a slightly different story, one which suggests that these situations are not so limited. Of the 99 opinions which addressed facts in which the defendant took the entirety of the plaintiff’s work, 27.3% found fair use (albeit with 9 of these 27 opinions finding a transformative use, and four finding a non-transformative use). The story is

²⁰¹ As between the creative/factual work inquiry and the published/unpublished work inquiry, which inquiry had the greater impact on the outcome of factor two? The data are inconclusive. All of the eleven opinions that addressed the use of a creative, unpublished work found that factor two disfavored fair use, and six of the seven opinions that addressed the use of a factual, published work found that factor two favored fair use. These results are not surprising. What happens, however, when the two subfactors under the second factor point in opposite directions? Twenty-two opinions addressed the use of a creative, published work and nineteen of these found that factor two disfavored fair use, while six opinions addressed the use of a factual, unpublished work and four of these found that factor two disfavored fair use. From this, we can tentatively conclude that a work’s creative status trumps its published status, but that a work’s factual status does not necessarily trump its unpublished status.

²⁰² 17 U.S.C. § 107.

²⁰³ Nevertheless, ten opinions evaluated the extent of the taking as a proportion of the defendant’s work, and two evaluated the extent of the taking both as a proportion of the plaintiff’s work and as a proportion of the defendant’s work.

²⁰⁴ Infinity Broad. Corp. v. Kirkwood, 150 F.3d 104, 109 (2d Cir 1998) (*quoting* NIMMER).

²⁰⁵ NIMMER, *supra* note 137, at § 13.05

more extreme in situations where the court finds that the defendant did or did not take the “essence” or the “heart” of the plaintiff’s work. Courts explicitly found that the defendant took the heart of the plaintiff’s work in 37 opinions, and found fair use in 35 of these. Courts made the opposite finding, that the defendant did not take the heart of the plaintiff’s work, in 25 opinions, and found fair use in 23 of these.

Given the apparent dominance of the first and fourth factors, it is easy to underestimate the importance of the third factor analysis to the outcome of the fair use test. But if we accept, as I will argue in a moment, that in nearly all situations, the fourth factor functions as a kind of metafactor under which courts synthesize their analyses of the first three factors, then we are in a position to appreciate how much of an impact the third factor actually has on the outcome of the fair use test—that is, on the outcome of the fourth factor. Here, the regression results reported in Table 9 are of special interest. They show that the subfactor considerations going to the taking of the entirety and in particular the heart of the work exert a significant influence on the outcome of the test. Specifically, the regression analysis predicts that if a defendant makes a commercial use of a creative, published work, its chances of succeeding in its fair use defense decline from 35.5% to 12.0% if it is found to take the entirety of that work, and to 1.0% if it is found to take the heart of that work.

D. Factor Four: Effect on the Market

The fourth factor calls upon courts to consider “the effect of the use upon the potential market for or value of the copyrighted work.”²⁰⁶ As mentioned above, the *Harper & Row* Court declared that this factor was “undoubtedly the single most important element of fair use,”²⁰⁷ and 59.0% of the opinions following *Harper & Row* (but preceding *Campbell*) explicitly cited this proposition. A decade later, the *Campbell* Court attempted to bend the *Harper & Row* dictum. Again choosing to reconstrue rather than rescind and replace, the Court obliquely stated that, in determining the question of fair use, “[a]ll [factors] are to be explored, and the results weighed together, in light of the purposes of copyright.”²⁰⁸ This intervention had a modest effect on the lower courts, as Figure 8 shows. Of the opinions following *Campbell*, 26.5% continued explicitly to state that factor four was the most important factor.

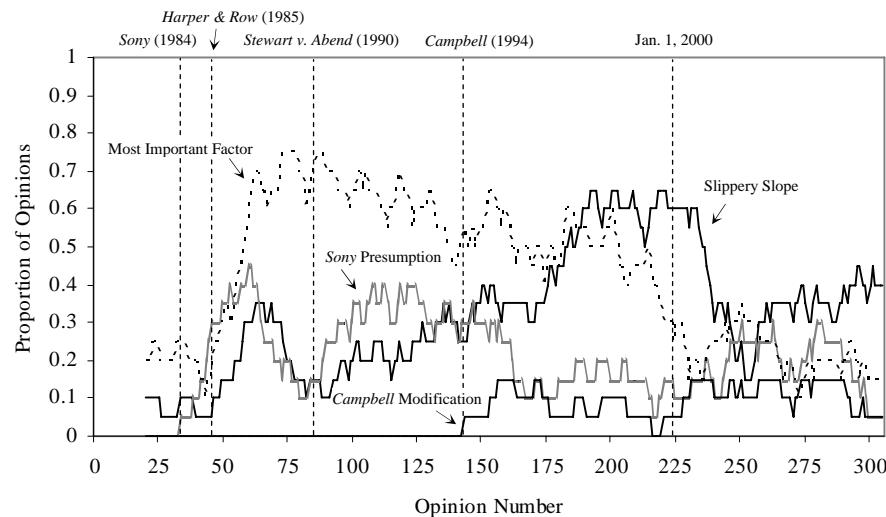
Whether the Supreme Court in its dicta on the issue was trying to describe the current state of the doctrine or prescribe what the doctrine should be is unclear. The conventional wisdom is that regardless of what the Supreme Court has said, the fourth factor analysis remains the most influential on the

²⁰⁶ 17 U.S.C. § 107.

²⁰⁷ *Harper & Row Publishers, Inc. v. Nation Enterprises*, 471 U.S. 539, 566 (1985).

²⁰⁸ *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 578 (1994).

FIGURE 8
20-OPIINION MOVING AVERAGE OF THE PROPORTION OF OPINIONS CITING VARIOUS FOURTH FACTOR PROPOSITIONS OF LAW



outcome of the overall test. The data support a different account, however, one which suggests that we have failed to appreciate the true role of the fourth factor analysis in the Section 107 test as applied. The fourth factor essentially constituted a metafactor under which courts integrated their analyses of the other three factors and, in doing so, arrived at the outcome not simply of the fourth factor, but of the overall test. Consider that of the 141 opinions that found that factor four disfavored fair use, 140 found no fair use. The one outlying opinion that found fair use reasoned, quite obtusely, that factor four “slightly disfavors a finding of fair use since plaintiff’s and defendant’s books are both directed towards the same audience.”²⁰⁹ Meanwhile, of the 116 opinions that found that factor four favored fair use, all but six found fair use. Four of the outliers reasoned that because there was no market for the plaintiffs’ works the factor could not favor the plaintiff,²¹⁰ and the other two were Scientology cases in which the court reasoned that the defendants’ limited uses did not substantially impact the religion’s efforts to retain or gain members.²¹¹

The synthetic and dispositive nature of the fourth factor analysis may explain why there have developed no real subfactors under factor four—other

²⁰⁹ Williamson v. Pearson Educ., No. 00-8420, 2001 U.S. Dist. LEXIS 17062, at *19 (S.D.N.Y. 2001).

²¹⁰ See Lamb v. Starks, 949 F. Supp. 753, 757 (N.D. Cal. 1996); Lish v. Harper’s Magazine Found., 807 F. Supp. 1090, 1104 (S.D.N.Y. 1992); Batesville Servs. v. Funeral Depot, Inc., No. 02-01011, 2004 U.S. Dist. LEXIS 24336, at *25 (S.D. In. 2004); Softel, Inc. v. Dragon Medical & Scientific Communs., No. 87-0167, 1992 U.S. Dist. LEXIS 9502, at *54 (S.D.N.Y. 1992).

²¹¹ See Religious Tech. Ctr. v. Netcom On-Line Commun. Servs., 923 F. Supp. 1231, 1248-49 (N.D. Cal. 1995); Religious Tech. Ctr. v. Lerma, No. 95-1107, 1996 U.S. Dist. LEXIS 15454, at *28-31 (E.D. Va. 1996).

than the three Section 107 factors that precede it—and why courts very rarely made specific factual findings under the factor. Instead, the vast majority of the opinions simply conducted what amounted to little more than an unstructured and conclusory rule of reason analysis.²¹² To the extent that courts structured their factor four analysis around anything, they structured it around various relatively feeble propositions of law. One is the presumption established by *Sony* that “[i]f the intended use is for commercial gain,” then a likelihood of market harm under factor four “may be presumed.”²¹³ The *Campbell* Court sought to overwrite this dictum on which the Sixth Circuit majority opinion in *Campbell* had so heavily relied. The *Campbell* Court first simply rejected the presumption outright as “a presumption which as applied here we hold to be error.”²¹⁴ Then the Court formulated a highly sensible refinement of the original presumption:

No “presumption” or inference of market harm that might find support in *Sony* is applicable to a case involving something beyond mere duplication for commercial purposes. *Sony*’s discussion of a presumption contrasts a context of verbatim copying of the original in its entirety for commercial purposes, with the noncommercial context of *Sony* itself (home copying of television programming). In the former circumstances, what *Sony* said simply makes common sense: when a commercial use amounts to mere duplication of the entirety of an original, it clearly supersedes the objects of the original and serves as a market replacement for it, making it likely that cognizable market harm to the original will occur. But when, on the contrary, the second use is transformative, market substitution is at least less certain, and market harm may not be so readily inferred.²¹⁵

But as Figure 8 shows, in the years following *Campbell*, many lower courts ignored or were ignorant of this refinement. Of the 108 opinions following *Sony* and preceding *Campbell*, 33 cited to *Sony*’s market harm presumption. Of the 161 opinions following *Campbell*, 25 continued to cite to the original presumption (with 76.0% of these ultimately finding no fair use), while only 14 cited to the *Campbell* modification of it (with 57.1% of these ultimately finding no fair use).²¹⁶

²¹² Cf. Leval, *supra* note 92, at 1107 (“Decisions are not governed by consistent principles, but seem rather to result from intuitive reactions to individual fact patterns.”).

²¹³ *Sony Corp. of America v. Universal City Studios, Inc.*, 464 U.S. 417, 451 (1984).

²¹⁴ *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 591 (1994).

²¹⁵ *Sony*, 464 U.S. at 451 (“A challenge to a noncommercial use of a copyrighted work requires proof either that the particular use is harmful, or that if it should become widespread, it would adversely affect the potential market for the copyrighted work.”). Though this dictum was directed towards noncommercial uses, courts, such as the *Campbell* Court, applied it, a fortiori, to commercial uses.

²¹⁶ Consider the remarkable post-*Campbell* Ninth Circuit opinion that cited to *Harper & Row* for the proposition that the fourth factor is the most important and then cited both to the *Campbell* modification and to the original *Sony* presumption. See *Elvis Presley Enters. v. Passport Video*, 349 F.3d 622, 631 (9th Cir.

Sony established another proposition of law, essentially a slippery slope principle, around which courts tended to structure their fourth factor analysis.²¹⁷ *Campbell* reinforced this proposition, stating that the fourth factor “requires courts to consider not only the extent of market harm caused by the particular actions of the alleged infringer, but also whether unrestricted and widespread conduct of the sort engaged in by the defendant would result in a substantially adverse impact on the potential market for the original.”²¹⁸ Figure 8 shows the impact on the fair use case law first of *Sony*’s and then of *Campbell*’s endorsement of this principle. Like many other doctrinal statements made by the Court, the slippery slope principle enjoyed an initial rapid rise in fashion after each of its Supreme Court endorsements and then suffered just as rapid a fall, until it reached a kind of equilibrium where it has recently been cited in about 35% of lower courts’ fair use opinions. The data suggest that there was no significant relation at any time between a court’s citing to the slippery slope principle and its finding of no fair use.

Ultimately, the paradox of the fourth factor is that it is everything in the fair use test and thus nothing. To assert, as a descriptive matter, that it is the most important factor—or, as a normative matter, that it is too important²¹⁹—is meaningless, primarily because it is no factor, no independent variable, at all. Instead, in practice, regardless of what we might hope—or fear—it would be, the actual doctrine of the fourth factor appears to consist of a few almost last-minute propositions of law that judges should keep in mind as they synthesize the various factual findings that they have made under the previous factors—and 93.6% of the opinions analyzed the factors in their numerical order. In theory, Section 107 instructs courts to conduct a balancing test along the four dimensions of its factors, “in light,” explains *Campbell*, “of the purpose of copyright.”²²⁰ In practice, judges appear to apply Section 107 in the form of a cognitively more familiar two-sided balancing test in which they weigh the strength of the defendant’s justification for its use, as that justification has been developed in the first three factors, against the impact of that use on the

2003) (“The last, and ‘undoubtedly the single most important’ of all the factors, is the effect the use will have on the potential market for and value of the copyrighted works. *Harper & Row*, 471 U.S. at 566. . . . The more transformative the new work, the less likely the new work’s use of copyrighted materials will affect the market for the materials. See *CBS Broad.*, 305 F.3d at 941. Finally, if the purpose of the new work is commercial in nature, ‘the likelihood [of market harm] may be presumed.’ *A&M Records*, 239 F.3d at 1016 (quoting *Sony*, 464 U.S. at 451).”).

²¹⁷ *Campbell*, 510 U.S. at 591.

²¹⁸ *Id.* at 590 (citations omitted).

²¹⁹ Leval, *supra* note 92, at 1124 (“Although the market factor is significant, the Supreme Court has somewhat overstated its importance. When the secondary use does substantially interfere with the market for the copyrighted work, as was the case in *Nation*, this factor powerfully opposes a finding of fair use. But the inverse does not follow. The fact that the secondary use does not harm the market for the original gives no assurance that the secondary use is justified. Thus, notwithstanding the importance of the market factor, especially when the market is impaired by the secondary use, it should not overshadow the requirement of justification under the first factor, without which there can be no fair use.”).

²²⁰ *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 578 (1994).

incentives of the plaintiff.²²¹ Factor four provides the analytical space for this balancing test to occur, and the various doctrinal propositions under factor four are merely there to tilt the scales one way or the other. In essence, like the four factors themselves, they are not legal propositions, but policy propositions.

V. CONCLUSION

At a general level, this Article has sought to show that the leading cases approach to our fair use case law is fundamentally flawed both as a descriptive and as a prescriptive enterprise. It is flawed as a descriptive enterprise in that it fails accurately to represent the actual state of our fair use doctrine as that doctrine is applied in the courts. Its anecdotal method is unequal to the task of accounting for a complex, multifarious system consisting of a multitude of judges, cases, and lines of doctrine. More importantly, but more subtly, the leading cases method is flawed as a prescriptive enterprise in that it perpetuates, because it requires, the myth that non-leading cases follow the lead of the leading cases. It is one thing to learn that academic scholarship has little apparent impact on the mass of our fair use case law,²²² but it is something else entirely to learn that Supreme Court and renowned circuit court opinions did not necessarily have much of an impact either. We cannot hope to better the true state of our fair use doctrine—the doctrine as it is practiced in the courts—if we continue to assume that conventionally agreed upon leading cases carry prescriptive force.

At a more specific level, this Article has sought to induce, albeit selectively and critically, from the population of opinions that it has studied a set of salutary doctrinal practices in the area of fair use. If this empirical precedent deserves any degree of deference, if it should itself carry any prescriptive force, it is not because it comes from a higher court, but because it represents what most judges from the population of judges who have applied fair use doctrine have seen fit to do. The assumption that the Article has made in this regard, and certainly one that is open to criticism, is that we may trust a population of judges over time systematically to point the way to the better practice of the doctrine. It appears that for all of their fractiousness, judges applying fair use doctrine have done just that. Where the non-leading cases declined to follow

²²¹ See generally Wendy J. Gordon, *Excuse and Justification in the Law of Fair Use: Commodification and Market Perspectives* (working paper on file with author).

²²² One rough index of the impact of academic scholarship on the fair use case law is the proportion of opinions that cited a law review article of any kind through the course of their fair use analysis. Of the seven Supreme Court opinions, four (or 57.1%) did so, while 25.0% of the 88 circuit court opinions and 10.9% of the district court opinions did so. However, if we exclude Judge Leval's *Towards a Fair Use Standard* article from this tabulation, the percentages decline to 50.0% of the Supreme Court opinions, 14.7% (or 13 of 88) of the circuit court opinions, and 5.7% (or 12 of 211) of the district court opinions. The law review article other than Leval's cited by the most courts was Gordon, *supra* note 9. Six courts cited to it.

the leading cases, they repeatedly—and systematically—did so in ways that expanded the scope of the fair use defense. To be sure, the data reveal many popular practices that impair the doctrine: courts have tended to apply the factors mechanically; they have sometimes made cynical use of the conflicting precedent available to them. These are systematic failures that require intervention. Nevertheless, though not each one on its own, but rather as a whole, the mass of non-leading cases have shown themselves to be altogether worthy of being followed.

APPENDIX: THE COLLECTION AND CODING OF THE OPINIONS

To collect a sample of the relevant opinions, I ran the following Lexis search in the Lexis *Federal Court Cases, Combined* database: copyright and “fair use” and 107 and date(geq (1/1/1978) and leq (12/31/2005)). This search yielded 557 opinions. I then ran the following Westlaw search in the Westlaw *All Federal Cases* database: copyright & "fair use" & 107 & da(aft 1977) & da(bef 2006). This yielded 575 opinions. A research assistant then reviewed the Lexis and Westlaw search results to identify any opinions reported uniquely in either Lexis or Westlaw. This review yielded a total 578 opinions from the two databases. A research assistant then reviewed each of these opinions to exclude those that did not involve in any way an issue of copyright fair use. This left 327 opinions. I then read each of these opinions and excluded an additional 20 opinions as irrelevant or only marginally relevant to copyright fair use. Of the 307 remaining opinions, 306 made substantial use of the Section 107 four-factor test, which I defined as any use of the test that made reference, however briefly, to at least two test factors. I then coded each of the opinions directly into an Excel 2003 SP2 spreadsheet according to a coding instrument consisting of 72 variables. I did the same in a new spreadsheet a second time and then compared the two spreadsheets for errors.

The coding instrument was designed to record (1) general data about the opinion (e.g., caption, citation, judge, venue, posture), (2) copyright-specific data about the opinion (e.g., the extent of the court’s treatment of the fair use defense, the disposition of the defense, whether the facts involved software, the reverse engineering of software, and/or the internet, whether the opinion addressed the First Amendment or parody), (3) factor-specific information about the opinion (e.g., which party the factor was found by the court to favor, how the court treated certain subfactor doctrine such as transformativeness or commerciality), and (4) various miscellaneous data about the opinion (e.g., whether the court cited any law review article, whether it cited legislative history, whether it relied on industry practice).

In a third round of coding, I loaded digital versions of the 306 opinions into Atlas.ti.²²³ I then conducted qualitative coding of each opinion according to a code list of 248 codes. One advantage of using Atlas.ti is that other researchers may review the Atlas.ti hermeneutic unit used for this study to verify each coding decision against the text of the opinions.

²²³ For other legal scholarship employing Atlas.ti, a qualitative data analysis program widely used in the social sciences, see, for example, Art Jipson, *The Post-September 11th Era: Interpretations of Security and Civil Liberties in the Political Margins of the Left and Right*, 2003 J. INST. JUST. INT'L STUD. 40; Ronald Weitzer, *Racialized Policing: Resident's Perceptions in Three Neighborhoods*, 34 LAW & SOC'Y REV. 129 (2000).

The coding instrument, Excel spreadsheet, Stata data file, and Atlas.ti hermeneutic unit are available upon request from the author and will be posted to the author's website upon publication of the Article.